Server and Management System Design for a Distributed VoD Service

A thesis submitted for the degree of Doctor of Philosophy in Electronic and Electrical Engineering

January 2000

Aikaterini Papanikolaou

UCL

Department of Electronic and Electrical Engineering
University College London

Statement of originality

The work presented in this thesis was carried out by the candidate. It has not been presented previously for any degree, nor is it at present under consideration by any other degree awarding body.

Candidate:

Aikaterini Papanikolaou

Diredtor of studies:

Dr Phil Lane

Statement of Availability

1. I authorise that the thesis presented by me in 2000 for examination for the MPhil./PhD degree of the University of London, shall, if a degree is awarded, be deposited in the University of London Library and in the library of the appropriate College and that, subject to the conditions set out in paragraph 9.2 (d) below, my thesis be made available for public reference, inter-library loan and copying.

2. I authorise the University of London to supply a copy of the abstract of my thesis for inclusion in any published list of theses offered for higher degrees in British universities or in any supplement thereto, or for consultation in any central file of abstracts of such theses.

3. I authorise the University of London to make a microfilm copy of my thesis for the purposes of inter-library loan and the supply of copies.

4. I understand that before my thesis is made available for public reference, inter-library loan and copying, the following statement will have been included at the beginning of my thesis: The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

5. I authorise the University of London to make a microfilm copy of my thesis in due course as the archival copy for permanent retention in substitution for the original copy. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

6. I understand that in the event of my thesis being not approved by the examiners, this declaration would become void.

Candidate:

Aikaterini Papanikolaou

Abstract

In this thesis a study on the design and performance of a distributed VoD service is presented.

In the first chapter a general introduction to the thesis is provided.

In the second chapter we provide a review to the enabling technologies for multimedia services and specifically for VoD. We review network, transport and delivery standards.

In the third chapter we present an overview of the most important VoD trials and an analysis on the composing components of the VoD service.

In the fourth chapter we present VoD service designs from the bibliography, that support and promotethe idea a distributed service provision system.

In the fifth chapter we present our design for a VoD server, supporting mass and interactive user connections, based on a two-tier control system. The first tier is based on staggered multicast at very short time intervals and processes normal connection flow and the second tier, provides the control for the user defined actions.

In the sixth chapter we present our design for the VoD server management system. Our system is a heuristics based one, it uses as little network status information as possible and makes decisions based on the current popularity rates it is receiving from the network. The movies are separated into classes depending on popularity and treated accordingly.

In the seventh chapter we present the simulation network we used in our study. The servers are placed in a regular square grid network, the movie library size is standard and the server capacity is varied.

In the eighth chapter we present the results we obtained using our simulation framework and applying different system and control parameters.

In the ninth chapter we compare our results of the distributed system to a

centralised system and to a system of non-networked independent servers.

In the last chapter the general conclusions are drawn.

Acknowledgements

I would like to express my gratitide to Dr Phil Lane for his guidance and support throughout the course of this study and the preparation of this thesis. I would also like to thank him for his friendship and moral support, for which I am deeply indebted.

I would also like to thank Professor John J. O'Reilly for his helpful remarks.

I wish to thank my first supervisor Dr Mark Wilby for his encourangement and support.

I would also like to thank a number of people for their controbutions, discussions and encouragement. A special "thanks" goes to Chris Matrakidis for his help.

I would also like to thank my friends in UCL Telecoms Group, for making these last years so enjoyable.

Last, but not least, I would like to thank my family, especially my mother and my husband for their support throughout all these years.

Financial support for this research has been provided by BT Laboratories Martlesham and EPSERC.

To my father, mother and Chari

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Multimedia service design and implementation has only recently been a topic of research as new technologies made their realisation feasible. VoD (Video-on-Demand) is only one of a range of interactive services; others include NoD (News-on-Demand), audio on demand, interactive shopping etc.

By definition VoD is the on demand delivery of films to the end user, with interactive functionality provided.

The main characteristics of the VoD service are:

1. support of multiple non-synchronous connections to one film
2. support of interactive functions on the delivered material
3. independent viewing for each user

These characteristics apply to VoD, and to interactive multimedia services in general, and they impose new requirements on the service provider in order to realise them.

This thesis presents a generic service design for VoD that will make it available to a large customer base. We have chosen a distributed VoD server model. The distributed model was chosen because it can offer scalability to cover future service demand trends, disperse material where is needed in the user network and can tolerate and adapt to partial server failure.

We identify the most important components of the service as being the following:

- the VoD server and
- the service management system

We propose a generic design for a scalable, interactive VoD server and a simple service management system.

The most important VoD server design requirements are:

- support of multiple connections to one film
- support of non-synchronised user connections to a number of films
- provision of real-time interactive actions such as pause and forward and backward scanning at any speed, or at random.
- scalable design to accommodate growing user demand

For the servers to operate successfully, a service control protocol is needed. The control protocol will manage movie and connection placement. The most important system resources have been identified as server space and network bandwidth. The control system is constructed in order to manage these two resources in the best possible way, while offering high film availability to the end user. We propose a simple management system based on the incoming user requests. Movies are classified according to these requests and treated accordingly. Furthermore, two system procedures are performed to improve film availability.

1.1 Organisation of thesis

The thesis is organised in ten chapters. The first is this introductory chapter.

In the second chapter we provide a review of the enabling technologies for multimedia services, and specifically for VoD. VoD is the interactive provision to the user of video material, such as movies, documentaries, video clips, etc. The provision of analogue video interactively to the end user is not profitable due to the large bandwidth required for the provision of one analogue channel (8MHz). Progress though on digital compressed video has led to the transmission of acceptable image quality at a rate as low as 1.5Mbit/s. This bit rate though far exceeds the rate that can be supported by twisted copper pair telephony connections that reach almost each home in this country and the EU. The appearance and progress in xDSL (Digital Subscriber Line) technology has brought the realisation of multimedia technology closer. xDSL technology is the easy alternative to a complete last loop upgrade to fibre, and is preferred by network providers because of its reduced installation cost. Additionally, CaTV coax cable networks provide extra bandwidth for this purpose. SDH and ATM and the upgrade of backbone networks to Tbits/s data rates provide the framework for a large scale implementation of distributed services. The next IP version, the IPng will provide extra functionality and support for time sensitive applications, will bring even closer the wide scale implementation of multimedia services.

In the third chapter we present an overview on the most important VoD trials performed by telecommunications operators, and introduce the most important VoD service elements.

In the forth chapter we compare the VoD service designs presented in the literature. These designs vary in concept and degree of interactive service provision. They either provide evidence to illustrate the benefits of a distributed solution compared to a centralised, or study a distributed VoD service network.

In the fifth chapter we present our design for a VoD server, supporting mass and interactive user connections. As the server is one of the most important elements of the VoD service, its performance will be reflected in overall service performance. Our server is based on a two tier architecture. On the first tier, staggered multicast at very short time intervals are used to provide the movie material, and normal flow control is performed. In the second tier the extra control for the interactive service is provided.

In the sixth chapter we present our design for the VoD server management system. As our VoD service design is a distributed one, the management system will be responsible for service operation. Our system is based on heuristics, it uses as little network information as possible and makes decisions based on the current popularity of movies received from the network. The movies are divided into three groups depending on popularity. The distance to the remote connections allowed for these movies is based on their popularity group. The servers cooperate in order to start a new movie and a search procedure is developed. Further, a relocation process is developed, which relocates less popular movies in order to make space available for more popular ones.

In the seventh chapter we present the simulation network we used in our study. The servers are placed in a regular square grid network, the movie library holds 100 movies and the server capacity is varied from 1 to 15 movies simultaneously available. The management scheme connects the users to their local server firstly, then if this is not possible an attempt is made to find the movie in a small cluster of neighbouring servers. If this fails the movie is started locally or remotely depending on network conditions. Relocation is allowed when connection is not possible and if relocation fails, a movie is discontinued, unavailable only when certain conditions apply in order to maintain QoS.

In the eighth chapter we present the results we obtained using our simulation framework and applying different system and control parameters. A changing probability profile is used to simulate change in demand during the course of a day. Finally, two different search ranges are studied for the relocation process, one where relocation is allowed to the whole of the network and one where relocation is limited to half the network.

In the ninth chapter we compare our results of the distributed system to a centralised system and to a system of independent servers. Our system performs better in all cases, and provides the best quality of service.

In the last chapter the general conclusions are drawn.

1.2 Summary of main contributions

The study presented in this thesis offers the design guidelines for the implementation of a real VoD service. The major contributions into four areas and can be summarised as follows:

- Design of an interactive, scalable multimedia server, where the design introduced is based on existing off the shelf components with added control.
- Design of a low complexity service control protocol which can be easily implemented.
- Service system requirements and network dimensioning in order to successfully support a VoD service.
- Least system settings required to satisfy system cost and QoS requirements.

The work carried out for this thesis has led to the following publications:

1. K. Papanikolaou and M. Wilby: "Design outline for Video–on–Demand server", *Fifth IEE Conference on Telecommunications*, Brighton UK 1995.
2. K. Papanikolaou and M. Wilby: "Administration System for Home Services offered by Metropolitan Area Networks", *Eighth IEEE Workshop on Local and Metropolitan Area Networks*, Berlin, Germany 1996.
3. K. Papanikolaou and M. Wilby: "Network and Control requirements for a distributed VoD Service", *IEEE Symposium on planning and design of broadband networks*, Montebello, Quebec, Canada, Oct. 96.
4. K. Papanikolaou and M. Wilby: "Architecture for a Networked VoD Service in a Multimedia Environment", *IEEE ROC&C/ITS'96*, Mexico 1996.
5. K. Papanikolaou and M. Wilby: "Control Protocol for Distributed Multimedia System Modelled on VoD", *IEEE Singapore International Conference on Communication Systems ICCS/ISPACS'96*, Singapore Nov.'96.

6. K. Papanikolaou and M. Wilby: "Management Scheme for VoD Server Network ", *Network and Optical Communications 1996: ATM Networks and LAN's, NOC'96*, Heidelberg 1996.

1.3 Summary

This chapter provides an introduction to the subject of the thesis; the design of a distributed VoD service. The main characteristics of both server and management system design are presented. The organisation of the thesis follows. Finally, the main contributions of the thesis are summarised along with a listing of the published papers based on the research work presented in this thesis. In the next chapter we present an overview of the multimedia enabling technologies.

Chapter 2

Supporting Technologies

The implementation of interactive multimedia services in the residential market has been brought closer due to a number of technological advancements that appeared at the beginning of the last decade [5]. These achievements cover a broad spectrum of fields, including image compression and transmission, processing power, memory and backbone networks. Highly efficient digitised video compression can considerably reduce the bandwidth requirements for acceptable video quality. Digital transmission and signal processing techniques can be used to increase the capacity of the standard twisted pair wire. Advancements in processing power and memory have also made possible the efficient handling of moving image data. Finally, the introduction of ATM based backbone networks, in conjunction with the other advancements in transport protocols such as IPv6, further facilitate the introduction of new services.

2.1 Introduction

In this chapter we present the most important of the VoD supporting technologies, which will play a significant role in the telecommunications and service networks of the future. First we present the image compression standards. These are the MPEG and JPEG digital image compression standards for moving and still images respectively. MPEG compresses the analogue moving picture signal from 150Mbit/s down to 1.5-40Mbit/s depending on quality. This has made video delivery to the home much more feasible than ever before.

Next we present an overview of the Access Networks of the future. Fibre is the only future proof medium, but the civil engineering costs make the network operators rather reluctant to implement it at this time. Copper pair is not yet obsolete thanks to ADSL, an xDSL digital transmission technology which can multiply copper pair bandwidth so that 2 to 3 Mbit/s are feasible for the home user. Coax networks mainly installed by CaTV operators can be used especially when upgraded for two-way communications. Wireless transmission can provide access with minimum costs of installation and maintenance to densely populated areas.

In the section on network technology we present ATM and IPv6 as two transport protocols that could carry the interactive multimedia services. Although ISDN as a digital services network has been proposed for quite some time, its implementation has taken telecommunications companies too long. Installation and usage costs are still considered high in most European countries, the UK included.

2.2 Compression Standards

Image compression has brought nearer the date of multimedia service implementation. Using image compression, the existing infrastructure can be used to deliver video to the residential user. The quality of the delivered image will be comparable to that now broadcast or delivered on cable. Additionally, image quality can be tailored to specific network capabilities.

2.2.1 JPEG

JPEG is a standardised image compression mechanism. JPEG stands for Joint Photographic Experts Group, the original name of the committee that developed the standard. It is designed for compressing either full-colour (24bit) or grey-scale digital images of natural (real-world) scenes. It is not very efficient in handling black and white (one bit/pixel) images, nor does it handle motion picture compression. Nevertheless, it could be used for still image compression when this is needed, for example in menus etc. JPEG uses the same compression techniques as MPEG apart from those especially used for temporal redundancy. Again the low frequency information is preserved as much as possible whereas the high frequency details are sacrificed for compression.

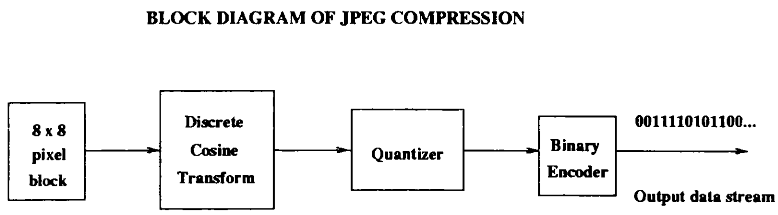

Figure 2.1: JPEG compression process

Figure 2.1 describes the JPEG process. JPEG divides the image into 8×8 pixel blocks, and then calculates the discrete cosine transform (DCT) of each block. A quantizer rounds off the DCT coefficients according to the quantization matrix. This step produces the "lossy" nature of JPEG, but allows for large compression ratios. JPEG's compression technique uses a variable length code on these coefficients. For decompression, JPEG recovers the quantized DCT coefficients from the compressed data stream, takes the inverse transforms and displays the image.

JPEG is lossy, in that the reconstructed image after decompression is not identical to the originally coded image. The algorithm achieves much of its compression by exploiting a known limitation of the human eye–it is less sensitive to high frequency details. Small colour details are not perceived as well as small details of light-and-dark. Thus, JPEG is intended for compressing images that are to be looked at by humans. When using JPEG compressed images in digital image processing we must be aware that they are different (detail is lost) to the original picture, due to the encoding process. A useful property of JPEG is that the degree of loss can be varied by adjusting the compression parameters. This allows the image maker to trade off file size against output image quality. Extremely small files can be made if high image quality is not required; this is useful for indexing image archives, making thumbnail views of icons etc. Conversely, high image quality can be achieved by relaxing the compression ratio.

Although it handles colour files well, it is limited in handling black-and-white and files with sharp edges (the compression ratio is rather small). The processing especially the compression cost, even on up-to-date computers, is also high. Further information is available from [6].

2.2.2 MPEG

Moving Pictures Expert Group (MPEG) is the name of the standard which has been produced by the ISO committee on digital colour video and audio compression. MPEG was initially developed for storage purposes. Provisions were therefore made in the algorithm to enable random access and fast-forward/reverse searches, when decoding from any digital storage media. However, the coding standard is flexible enough to be suitable for a much wider range of video applications. Appropriate modifications may be introduced in order to reduce the end-to-end delay due to the transmission process and further adapt the standard to network applications [7, 8].

MPEG is a highly efficient algorithm and can produce up to a 200:1 compression ratio for fairly unchanging video sequences. The compression rate depends on the video being coded, fast changing scenes dictate a lower compression ratio. In MPEG both spatial and temporal information redundancy is removed. The MPEG standard has three parts: the video encoding, the audio encoding, and the systems part. In the video encoding part, spatial and temporal redundancies are removed; only the frames without any such redundancy are coded to full extent. Audio is also coded to reduce bandwidth requirements. The system part includes synchronisation information of the audio and video streams which is going to be used during reconstruction.

Two features of the algorithm are the most important from the video viewpoint. These are the coding layers and the coding modes. The video frame is hierarchically segmented into coding layers. By segmenting the frame the particular features of the algorithm can be used to its full extent to improve compression efficiency. Highly detailed areas are coded using more data, thus keeping image quality at acceptable levels, whereas simpler areas can be coded more efficiently.

The coding modes are determined by the information in each frame. Frames which designate a scene change are unique and therefore coded to a full extent without any reference to previous or next frames. The frames between two scene changes are coded according to their resemblance. Some frames are only dependent on previous frames, but others are dependent on previous as well as following frames. Accordingly, the compressed frames may not be broadcast in absolute time sequence.

Layers and modes

The video sequence to be MPEG compressed is usually divided into four layers, each with different use and functionality. The layers are arranged as follows.

- **Block**-A block is the smallest coding unit in the MPEG algorithm. It is made up of 8×8 pixels and can be of three types: Luminance (Y), red chrominance (Cr) and blue chrominance (Cb). The block is the basic unit for intraframe DCT (Digital Cosine Transform) coded blocks.

- **Macroblock**- A macroblock (MB) is the basic coding unit in the MPEG algorithm. A macroblock is a 16×16 pixel segment of Luminance components and the corresponding 8×8 pixel section of the two chrominance components. Because the human eye is not very sensitive to the chroma region changes, as compared to the luminance region, the chroma matrices are usually decimated or reduced in size by a factor of two in both the vertical and the horizontal directions. Consequently there are one forth the number of chrominance pixels to process as there are luminance pixels. This format, referred as 4:2:0 format, is employed in MPEG-1. MPEG-2 additionally allows for either no decimation or only horizontal decimation of the chroma component. These two formats are referred to as 4:4:4 and 4:2:2 formats respectively. Pictures can be categorised into three main types based on their compassion schemes. Since each vertical component has one-half the vertical and horizontal resolution of the Luminance component, a macroblock consists of 4 Y, 1 Cr and 1 Cb block.

- **Slice**- A slice is a horizontal strip of macroblocks within a frame. Coding operations on blocks and macroblocks can only be performed when all pixels for a slice are available. A slice is an autonomous unit since coding for a slice is done independently from its neighbours.

  Slices are important in the handling of errors. If the bitstream contains an error, the decoder can skip to the start of the next slice. Having more slices in the bitstream allows better error concealment, but uses bits that could otherwise be used to improve picture quality.

- **Picture**- A picture in MPEG terminology is the basic unit of display and corresponds to a single frame in a video sequence.

Figure 2.2 shows the MPEG layer hierarchy, from blocks to the frame sequence.

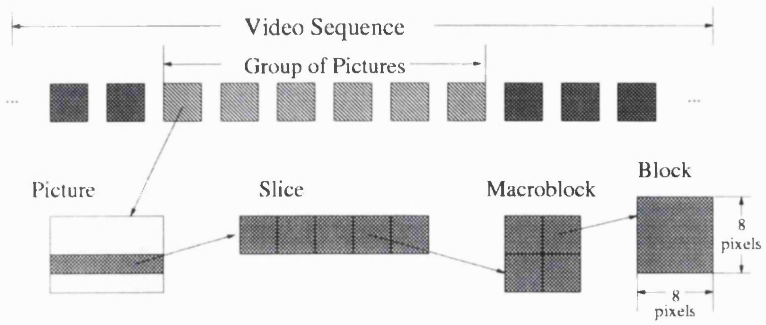

Figure 2.2: MPEG hierarchy

A picture consists of three rectangular matrices representing luminance (Y) and two chrominance (Cb and Cr) values. The Y matrix has an even number of rows and columns. The Cb and Cr matrices are one-half the size of the Y matrix in each direction (horizontal and vertical).

The spatial dimensions of the frame are variable and determined by the application requirements. A frame made up of 485×720 pixels corresponds to the NTSC standard or 576×720 to PAL, a similar resolution in MPEG would be sufficient to replace the analogue standards. VHS requires an even lower resolution of 352×240.

In the MPEG coding scheme, a choice of several coding modes is available at the frame and macroblock level. For encoded frames three modes can be utilised:

- **I (intra)** frames are coded as still images.
- **P (forward predicted)** frames can be predicted from the most recent I or P frame.
- **B (bi-directionally predicted)** frames are interpolations between I and P frames.

Each frame is encoded according to its properties. Typically, the I frames are sent once every 10 or 12 frames. Reconstructing a B frame for display requires the preceding and following I and/or P frames. The coded frames are sent out of time-order, to compensate for the B-pictures which need the next picture present for reconstruction. In a network environment, this would require bigger buffers, so that all the required frames could be stored on a buffer before reconstruction. The system could be prone to delays and consequently loss of image quality if no guarantees are given on transfer time by the transport protocol.

In more detail, intra pictures, or I-pictures, are coded independently, using only information present in the picture itself. I-pictures provide potential random access points into the compressed video data. The whole I frame undergoes 8×8 block-based Discrete Cosine Transform (DCT) for the luminance and chrominance components, without referring to other frames. The DCT coefficients are then quantized. The DCT coefficients of individual blocks are coded differently within a slice. I-pictures typically use about two bits per coded pixel.

Predicted pictures, are coded with respect to the nearest previous I or P picture. This technique is called forward prediction. As for I pictures, P pictures serve as a prediction reference for B pictures and future P pictures. However, P pictures use motion compensation (MC) which is discussed in section 2.2.2 to provide more compression than is possible with I pictures. For interframe-coded frames (P), temporal redundancy is first reduced by causal MB (Macro-Block)-based motion compensation, a temporal redundancy eliminating technique, with respect to the preceding I or P frames stored in the Frame Storage. If the motion estimation (ME) error is less than a threshold (i.e. there is enough redundancy the interframe-coding is worthwhile), then the motion vector (MV) will be differentially coded using variable length coding (VLC), while the ME error will undergo DCT, coarse quantization, and then VLC. Otherwise, that MB will undergo intraframe-coding. Differential coding is used because it reduces the total bit requirement by transmitting the difference of the motion vectors of consecutive frames. The compression efficiency and the quality of the reconstructed video depend on the accuracy of the motion estimation. Unlike I pictures, P pictures can propagate coding errors because P pictures are predicted from previous reference (I or P) pictures.

Bidirectional pictures, or B pictures, are pictures that use both a past and future picture as a reference for motion compensation. This technique is called bidirectional prediction. *Figure 2.3* illustrates the relation of the three picture types. Prediction is non-causal since both past and future frames are used. B pictures provide the most compression and do not propagate errors because they are never used as a reference, but display already propagated errors. Bidirectional prediction also decreases the effect of noise by averaging two pictures.

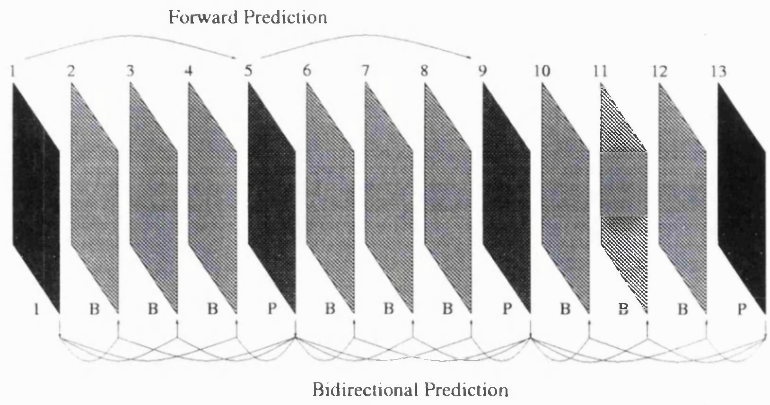

Figure 2.3: MPEG standard coding example

Bidirectionally predicted (B) frames are coded in a way similar to coding P frames; however, the motion compensation is bidirectional with respect to both preceding and following P (or I) frames [9].

An MPEG coded sequence is characterised by three parameters: q, N and M. The quantization factor q controls the degree of fitness of quantization. $N-1$ is the number of frames coded between successive I frames, while $M-1$ is the number of B frames coded between successive P frames. A group of frames (GOF) with $N=10$, $M=3$ is as follows:

IBBPBBPBBP

Data in an MPEG coded video stream are of unequal importance. When data in I or P frames are lost during transmission, the frame contents in the Frame Storage at the coder and decoder become different. Even if no further data is lost, for the following P or B frames, the ME at the coder and decoder will refer to different frame contents at the "baseline" of estimation. Consequently errors due to data loss of one I or P frame will propagate along the following P and B frames, and this is often referred to as error propagation. The accumulated errors can be cleared by sending an I frame.

MPEG Coding

The MPEG transform coding algorithm includes these steps:

- Discrete cosine transform (DCT)
- Quantization
- Run-length encoding

Both image blocks and prediction-error blocks have high spatial redundancy. To reduce this redundancy, the MPEG algorithm transforms 8×8 blocks of pixels or 8×8 blocks of error terms from the spatial domain to the frequency domain with the Discrete Cosine Transform (DCT). DCT is a time to frequency transform, which exploits the spatial correlation of the pixels by converting them to a set of independent coefficients. The low frequency coefficients contain more energy than the high frequency ones. This process allows the high energy, low frequency coefficients to be coded with greater number of bits, while using fewer or zero bits for the high frequency, low energy coefficients. Equation 2.1 represents the two-dimensional DCT transform operation that should be applied on a 8×8 pixel block. This is done separately for the luminance and chrominance components. The two-dimensional DCT can be separated in a number of sequential one-dimensional DCTs, for the 8×8 case, eight 8×1 DCTs followed by eight 1×8 DCTs. DCT is quite a complex computation and a number of algorithms have been developed to reduce its complexity [10], for example:

$$F(u,v) = \frac{1}{4}C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7}\cos(\pi(2x+1)u/16)\cos(\pi(2y+1)v/16) \tag{2.1}$$

where,

$$C(u), C(v) = \begin{cases} 1/\sqrt{2} & \text{for } u, v = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Next, the algorithm quantizes the frequency coefficients. Quantization is the process of approximating each frequency coefficient as one of a limited number of allowed values. The encoder chooses a quantization matrix that determines how each frequency coefficient in the 8×8 block is quantized. Human perception of quantization error is lower for high spatial frequencies, so high frequencies are typically quantized more coarsely (i.e., with fewer allowed values) than low frequencies.

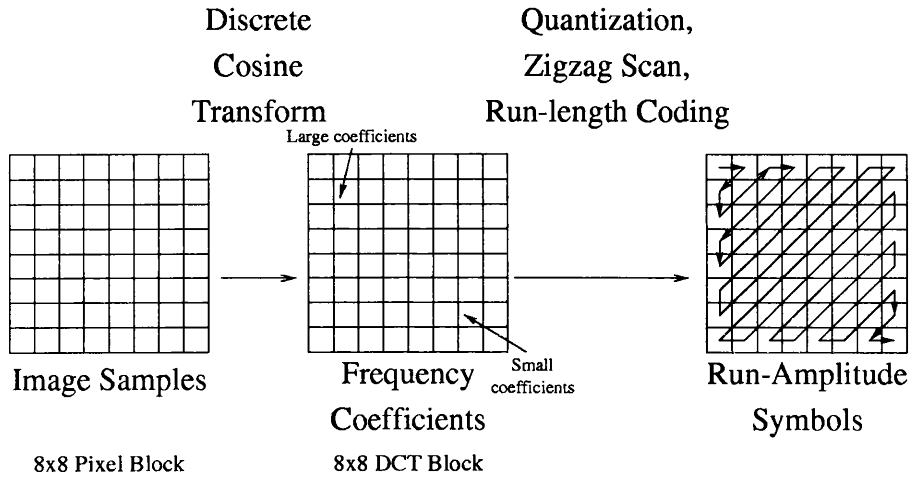

Figure 2.4: Transform Coding Operations

Some blocks of pixels need to be coded more accurately than others. For example, blocks with smooth intensity gradients need accurate coding to avoid visible block boundaries. To deal with this inequality between blocks, the MPEG algorithm allows the amount of quantization to be modified for each macroblock of pixels. This mechanism can also be used to provide smooth adaptation to a particular bit rate.

The transform exploits the spatial correlation of the pixels by converting them to a set of independent coefficients. The low frequency coefficients contain more energy than the high frequency ones. These coefficients are quantized employing a quantization matrix. This process allows the high energy, low frequency coefficients to be coded with a greater number of bits, while using fewer or zero bits for the high frequency, low energy coefficients. The high frequency coefficients can afford to be dropped because the eye lacks the ability to detect high frequency changes. Retaining only a subset of the coefficients reduces the total number of parameters needed for representation by a substantial amount. The process is identical for the luminance and the chrominance pixel blocks.

MPEG TOOLS	*REDUNDANCY EXPLOITED*
DCT	Spatial
MC Prediction/Interpolation	Temporal
Run Length/Huffman	Coding
Differential Coding	Temporal

Table 2.1: MPEG Tools

However, since the human visual sensitivity to the luminance and chroma varies, the quantization matrices for the two differ. The quantization process also helps in rate control, i.e. allowing the encoder to output bitstreams at a specified bitrate. The DCT are coded employing a combination of two lossless coding schemes-Run Length and Huffman. The coefficients are scanned in a zigzag pattern to create 1-D sequence (see *Figure 2.4*). MPEG-2 can additionally provide a different scan pattern as an alternative. The resulting 1-D sequence usually contains a large number of zeros due to the lowpass nature of the DCT spectrum and the quantization process. Each non-zero coefficient is associated with a pair of pointers. First, its position in the block which is indicated by the number of zeros between itself and the previous non-zero coefficient. Second, its coefficient value. Based on these two pointers, it is allotted a variable length code from a lookup table. This is done in a manner so that a highly probable combination gets a code with fewer bits, while the unlikely ones get longer codes. Adopting this lossless coding technique the total number of bits is kept down. However, since spatial redundancy is limited, the I pictures provide only moderate compression. These pictures provide important hooks for random access into the digital bitstream for editing and viewing purposes.

MPEG is a clever combination of a number of diverse tools that exploit temporal and spatial redundancy. It is efficient and can offer a variety of qualities according to the application. Advancements in coding and processing have made the realisation of real-time video services possible. Table 2.1 summarises the tools used.

Video Stream Composition

The MPEG algorithm allows the encoder to choose the frequency and location of I pictures. This choice is based on the application's need for random accessibility and the location of scene cuts in the video sequence. In applications where random access is important, I pictures are typically used two times a second.

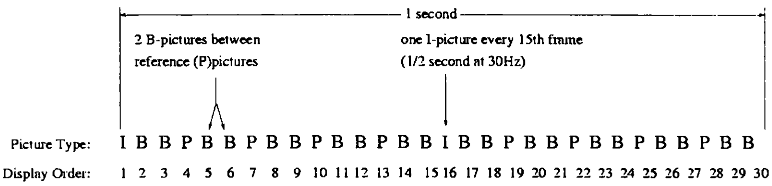

Figure 2.5: Typical display order of picture types

The encoder also chooses the number of B-pictures between any pair of reference (I or P) pictures. This choice is based on factors such as the amount of memory in the encoder and the characteristics of the material being coded. For example, a large class of scenes have two bidirectional pictures separating successive reference pictures. A typical arrangement of I, P, and B pictures is shown in *Figure 2.5* in the order in which they are displayed.

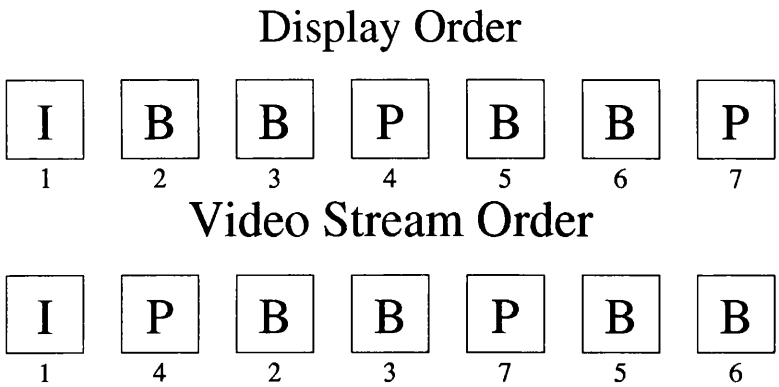

Figure 2.6: Typical display order of picture frames

The MPEG encoder reorders pictures in the video stream to present the pictures to the decoder in the most efficient sequence. In particular, the reference pictures needed to reconstruct B pictures are sent before the associated

B pictures. *Figure 2.6* demonstrates this ordering for the first section of the example shown above.

Motion Compensation

Motion compensation is a technique for enhancing the compression of P and B pictures by eliminating temporal redundancy. Motion compensation typically improves compression by about a factor of three compared to intra-picture coding. Motion compensation algorithms work at the macroblock level.

When a macroblock is compressed by motion compensation, the compressed file contains the following information:

- The spatial vector between the reference macroblock(s) and the macroblock being coded (motion vectors)
- The content differences between the reference macroblock(s) and the macroblock being coded (error terms)

Not all information in a picture can be predicted from a previous picture. Consider a scene in which a door opens: The visual details of the room behind the door cannot be predicted from a previous frame in which the door was closed. When a case such as this arises–i.e., a macroblock in a P picture cannot be efficiently represented by motion compensation–it is coded in the same way as a macroblock in an I picture using transform coding techniques.

The difference between B and P picture motion compensation is that macroblocks in a P picture use the previous reference (I or P picture) only, while macroblocks in a B picture are coded using any combination of a previous or future reference picture.

Four codings are therefore possible for each macroblock in a B picture:

- Intra coding: no motion compensation
- Forward prediction: the previous reference picture is used as a reference
- Backward prediction: the next picture is used as a reference

- Bidirectional prediction: two reference pictures are used, the previous reference picture and the next reference picture

Backward prediction can be used to predict uncovered areas that do not appear in previous pictures.

Audio Stream Data Hierarchy

The MPEG standard defines a hierarchy of data structures that accept, decode and produce digital audio output. The MPEG audio stream, like the MPEG video stream, consists of a series of packets. Each audio packet contains an audio packet header and one or more audio frames.

Each audio packet header contains the following information:

- Packet start code - Identifies the packet as being an audio packet.
- Packet length - Indicates the number of bytes in the audio packet.

An audio frame contains the following information:

- Audio frame header - Contains synchronisation, ID, bit rate, and sampling frequency information
- Error-checking code - Contains error-checking information
- Audio data - Contains information used to reconstruct the sampled audio data.
- Ancillary data - Contains user-defined data.

Audio and video mixing

When audio is digitally stored, like on a conventional audio CD, a method called Pulse Code Modulation (PCM), is usually used. Despite the fact that audio stored in this manner takes up a lot less space than regular image information, it still is not sufficiently compact compared to MPEG-Video. For this reason, the sound is also compressed to match a level of compactness that is proportional to the compressed image information. This audio compression technique used is MPEG-Audio compression according to ISO 11172-2. This method makes use of Huffman-encoding and the psycho-acoustic model. The Huffman-encoding exploits repetitive patterns and redundancy, while the psycho-acoustic model exploits the shortcomings of the human brain, which is incapable of simultaneously "hearing" different sound frequencies that are close together in the spectrum. The brain will only (or mainly) detect the loudest or most prominent frequency, and ignore almost all frequencies close around this prominent frequency that aren't loud enough. For this reason, it is not necessary to store all sound information, but only the "remarkable" parts. This method does not effect stereo or Dolby surround pro-logic information. The result is very compactly stored sound information, with near CD-perceived audio quality.

MPEG-2 audio is a compatible extension to MPEG-1 audio encoding, which enables the transmission of mono, stereo, or multichannel audio in a single bitstream. It can operate at a wide range of bit rates (8 kbit/s up to more than 1 Mbit/s) and supports sampling rates of 16, 22.05, 24, 32, 44.1 and 48 kHz. For stereo, a typical application would operate at an average bit rate of 128-256 kbit/s. A multichannel movie soundtrack requires an average bit rate of 320-640 kbit/s, depending on the number of channels and the complexity of the audio to be encoded. MPEG-2 defines an extension for five full bandwidth channels plus a low frequency enhancement (LFE) channel, termed 5.1 multichannel. With an additional compatible extension, seven channels are possible (7.1 multichannel).

In devising an encoding method, the basis had to be the human ear. Although not a perfect device for acoustic reception, advantage was taken of one of its characteristics: a non-linear and adaptive threshold of hearing. The threshold of hearing is the level below which a sound is not heard. It varies with frequency and, of course, between individuals. Most people's hearing is most sensitive between 2 and 5 kHz. Whether a person hears a sound or not depends on the frequency of the sound and whether the amplitude is above or below that person's hearing threshold at that frequency.

The threshold of hearing is adaptive, and is constantly changed by the sounds heard. For example, an ordinary conversation in a room is perfectly audible under normal conditions. However, the same conversation in the vicinity of a loud noise, such as an aircraft passing low overhead, is impossible to hear due to the distortions introduced to the hearing thresholds of the individuals concerned. When the aircraft has gone the hearing thresholds return to normal. Sounds that are inaudible due to dynamic adaptation of the hearing threshold are said to be "masked".

This effect is universal but is of particular relevance in music. An orchestra instrument playing fortissimo will, to a greater or lesser extent, make the sound of some other instruments inaudible to the human ear. When the music is recorded, however, all the frequencies go on the medium because the response of the recording device is flat, i.e. it is not dynamically adaptive. When the recording is played the masked instruments will not be audible to the listener. A linear recording, as used on CD, is inefficient in this respect. To make the best use of a recording medium the parts of the medium that contain inaudible data can better be used to store audible data. In this way a fixed capacity recording medium can contain a considerably increased amount of audio without any loss of quality, as is perceived by the human ear. Also, the demands on a transmission link carrying the information are reduced.

The MPEG-2 standard was designed with compatibility being a major consideration. With the ever growing number of applications of MPEG-1 audio, especially in the entertainment, satellite broadcasting and multi-media fields, this compatibility will provide the consumer a cross-platform format to enjoy high quality audio reproduction. The core of the MPEG-2 bitstream is an MPEG-1 bitstream which enables fully compatible decoding by an MPEG-1 audio decoder. In addition, the need to transfer two separate bit streams (one for stereo and another one for the multichannel audio program) is avoided. In other words, a future upgrade of e.g. DSS (Digital Subscriber System) with multi-channel audio will not make existing set top boxes obsolete. The existing ones will reproduce stereo, the new ones high quality multichannel sound.

An MPEG-1 decoder will be supplied in the MPEG-1 part of the bitstream with an appropriate (stereo) "downmix" of all channels in the multichannel frame. The left and right channel of the stereo signal contain components of all the channels, according to the equations in the compatibility matrix. The MPEG-1 (stereo) decoder decodes the stereo part of the MPEG-2 frame, and ignores the multichannel extensions. MPEG-2 defines four matrix sets, one of which is selected in the decoder from information in the MC (multichannel) frame header.

Synchronisation

The MPEG standard provides a timing mechanism, using time stamps that ensures synchronisation of audio and video. This method allows the provision of flexibility in decoder design, packet length, audio sample rates, coded data rates and network performance. The decoder can use the time stamp information to display frames accordingly, frames with expired time stamps can be discarded, whereas frames that have arrived early can be saved in a buffer. The standard includes two parameters: the system clock reference (SCR) and the presentation timestamp (PTS).

The MPEG-specified "system clock" runs at 90 kHz. System clock reference and presentation timestamp values are coded in MPEG bitstreams using 33 bits, which can represent any clock cycle in a 24-hour period.

For example C-Cube Microsystems uses a System Clock Reference (SCR) called CL480 to help synchronise picture and sound during reproduction.

An SCR is a snapshot of the encoder system clock which is placed into the system layer of the bitstream, as shown in *Figure 2.7.* During decoding, these values are used to update the system clock counter in the CL480.

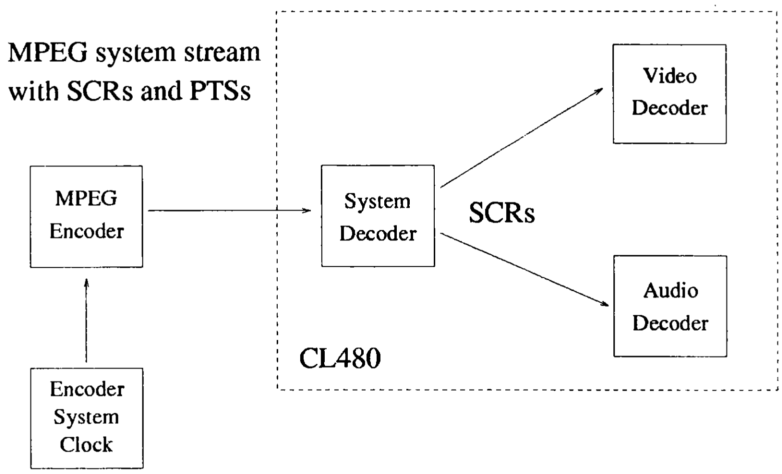

Figure 2.7: SCR Flow in MPEG System

Presentation Timestamps

Presentation timestamps (PTS) are samples of the encoder system clock that are associated with video or audio presentation units. A presentation unit is a decoded video picture or a decoded audio time sequence. The PTS represents the time at which the video picture is to be displayed or the starting playback time for the audio time sequence.

The decoder either skips or repeats picture displays to ensure that the PTS is within one picture's worth of 90 kHz clock ticks of the SCR when a picture is displayed. If the PTS is earlier (has a smaller value) than the current SCR, the decoder discards the picture. If the PTS is later (has a larger value) than the current SCR, the decoder repeats the display of the picture.

MPEG versions

There are four MPEG (two of these still developing) standards varying in bit-rate and in the quality of the moving picture offered. MPEG-1 was originally designed for delivery of video to consumer devices when single speed CD-ROM data rates were 1.5Mbit/s, supporting super VHS quality video.

MPEG-2 is a standard for digital television and offers better resolution and quality than MPEG-1, having been designed for delivery of broadcast and

HDTV quality video. MPEG-2 is already being used for Direct Satellite Broadcast and HDTV quality video.

MPEG-2 is designed to offer higher quality at a bandwidth of between 4 and 10 Mbit/s. This is achievable using today's CD-ROM technology. MPEG-2 will compress, for example a 720×480 full-notion video in broadcast television and video-on-demand applications. It has several advantages comparing to MPEG-1:

- compression ratio of interleaved picture
- transmission stream is suitable for packet switched computer networks
- supporting wide range of formats including the normal TV picture size and HDTV
- data streams of wide bandwidth can be accommodated

MPEG-2 over ATM has several problems. For example, the transmission delay, delay jitter and error correction are areas of intensive work today. The ATM forum SAA/AMS groups phase 1 specification for MPEG-2 over ATM for VoD service is evolving rapidly. It aims to use ATM equipment which is available today. The user should have to buy only a MPEG card, leaving it up to the equipment manufacturers to solve interoperability problems. First, the networked services are assumed to restrict the magnitude of proprietary solutions, and secondly what remains in the implementation side of the problem can be solved in the MPEG application.

An interoperability problem can rise when more than one incompatible versions of the same algorithm are applied in order to create proprietary systems which will not allow the customer to move freely from service provider to service provider. These types of solutions will restrict rather than promote the wider implementation of the proposed technology.

Furthermore, there are MPEG-4 and MPEG-7. MPEG-4 became an International standard from January 1999. It is based on three fields: digital television, interactive graphics applications (synthetic content) and the World Wide Web (distribution and access to content) and will provide standardised technological elements enabling the integration of production, distribution and content access paradigms in the three fields. It could also be used in videophones etc. Bit rates targeted for the MPEG-4 video standard are between 5 - 64kbit/s for mobile video applications and up to 2 Mbit/s for TV/film applications.

MPEG-7 is called Multimedia Content Descriptor Interface and its objective is to extend the incorporation of more data types and provide fast access to material via indexing and easy search capabilities. MPEG-7 will specify a standard set of descriptors that can be used to describe various types of multimedia information. For example when searching for a tune someone could whistle the tune. Material that has MPEG-7 data associated with it, can be indexed and searched for.

Substantial computing power is required to encode MPEG data varying with the required version and specified compression parameters, perhaps several hundred MIPS to encode 25 frames/s, though for non real-time compression as in VoD, this requirement can be relaxed. Decoding though, is not so demanding. A number of companies are active in producing MPEG encoding and decoding chips. As decoding requires less processing power this will make it quickly available to the end user STB (set-top-box). By the time the service is widely available, costs will drop to an acceptable level.

Real time MPEG-1 and -2 decoding already exists for the Personal Computer market, in the form of chips as well as software. Agreement on a common standard will facilitate further the spread of the new technology. MPEG decoding cards could be mass produced for the residential market and incorporated in the Customer Premises Equipment (CPE). Interoperability problems could arise–different proprietary versions of MPEG could require different home equipment and thus limit user choice. A global standard would further facilitate the cost-effective delivery of services dependent on the standard to the mass market.

2.3 Delivery transport network/Access transmission

The provision of the new services will demand at least a partial network upgrade. The core as well as the access network have to be upgraded in order to accommodate the new services. Fibre is known for its high bandwidth potential. It is already installed in many parts of the backbone network in Europe as well as in the US. Extending the use of fibre into the community network as well as to the local loop though, will demand considerable investment. Most telecommunications network providers are seeking an intermediate, cheaper solution which will enable the provision of multimedia services and thereafter finance further network upgrade.

2.3.1 Copper pair access

The great advantage of using the existing copper pair for VoD services is that copper pairs are already installed for telephony, and the provision of VoD can therefore be considered as a marginal technical and fiscal cost especially for the incumbent operator. Telecommunications regulators should allow the other telecommunications providers to use the existing infrastructure at a reasonable cost under the fair trading and competition principle. The unbundling of the local loop to new providers will be beneficial for the end user primarily since it will promote competition and give the end user a wider choice of services and providers.

The technique used for the transmission of 1.5Mbit/s to 6Mbit/s over copper pair is known as Asymmetrical Digital Subscriber Loop (ADSL) [11, 12] and has been recently standardised by a number of international standardisation committees such as ETSI, ANSI etc. ADSL belongs to a family of standards known as *x-DSL*, which will be used to provide symmetric, asymmetric high and low bandwidth services to the end user via the copper pair of the local loop.

2.3.2 ADSL

The Asymmetric Digital Subscriber Line (ADSL) system is an asymmetrical bi-directional transmission system used in the local subscriber loop between the local telephone switch and the subscriber's home, thus allowing the transmission of broadband services using the existing twisted cable pair [13]. ADSL though has a range limited to a few kilometres. Below this range signal regenerators are not needed.

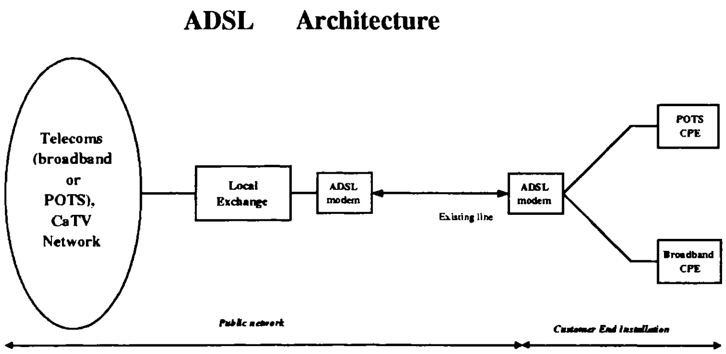

Figure 2.8: ADSL Architecture diagram

ADSL technology requires two modems: one at the customer premises and one at the local exchange, *see Figure 2.8.* It can support both POTS and high bandwidth multimedia services. It provides a 1.5 to 6 Mbit/s downstream channel and a 16 to 640 kbit/s upstream channel system depending on cable length. ADSL is only one of a family of digital subscriber lines [14]. Other variants include:

- **HDSL** - High data rate DSL, based on the idea of inverse multiplexing, bandwidth limited to 250 kHz by adoption of appropriate modulation techniques, span of 3.5 km and 2.048 Mbit/s of symmetric bandwidth.

- **SDSL** - Single line DSL, supports HDSL rates over a single line (2.048 Mbit/s of bandwidth)– compatible therefore with residential customers, range of about 3 km.

- **RADSL** - Rate adaptive DSL. Supports 32 kbit/s to 9 Mbit/s downstream and 32 kbit/s to 1.5 Mbit/s upstream. Span depends on the data rate.

- **VDSL** - Very high data rate DSL. It can support 12.96, 25.92 and 51.84 Mbit/s downstream and 2 to 20 Mbit/s upstream, the span varies from 1.5, 1 and 0.3 km. It requires though a FTTC system to support the VDSL bandwidth that is required by the application.

ADSL technology is the result of at least three advancements in the field of digital communications. First, ADSL takes advantage of the more general advancements in adaptive digital filtering and DSP applied to digital communications.

Techniques such as Trellis coding and forward error correction with interleaving are found in ADSL systems, to increase channel capacity in the presence of Gaussian noise and to provide immunity to impulsive noise respectively. New techniques, such as asymmetric echo cancellation, will find application for the first time in ADSL. Also, a relatively new modulation format, discrete multitone (DMT), has been selected by the American National Standards Institute (ANSI). A forum has been established called the ADSLF, for standardisation of ADSL systems.

Carrierless AM/PM (CAP), Discrete Multi-tone Transmission (DMT), and Frequency Division Multiplexing are to be considered as the modulation techniques. ADSL systems are used at local subscriber loops with telephone or basic ISDN access.

ADSL range is limited by a number of factors such as impulse noise and cross talk mainly. From practical experiments, it appears that cross talk is the most damaging one. Cross talk is inherent to inductive coupling between adjacent pairs due to its effect two types of cross talk have been defined: NEXT (near-end cross talk) and FEXT (far-end cross talk). The principle of NEXT is that a fraction of the power of the signal in one pair is detected in the reverse direction on another pair, belonging to the same distribution wire. In FEXT a fraction of the power sent by the local exchange on one pair is detected on another pair belonging to the same distribution wire.

Another advance making ADSL possible is continuing increases in the complexity and speed of very large scale integration (VLSI) technology. VLSI devices that implement the majority of ADSL transceivers will make the technology cost-effective, sufficiently so for a consumer electronics-type application like video dial tone to become feasible.

Finally, ADSL takes full advantage of a detailed understanding of the Local Exchange distribution network; that part of the network extending from the Local Exchange, where switching and trunking equipment is located, to the customer's home. The great bulk of this distribution network today consists of copper pairs providing normal telephone service to residences and businesses, but increasingly fibre/copper digital transmission systems are entering the plant. Fibre could extend to a RCU (remote control unit) site, in a vault, a hub or a cabinet, from there a copper pair network, which in future could incorporate fibre, extends to the local exchange. This portion of the network referred to as the "feeder", may also contain copper based trunking equipment in the form of E–carrier systems.

The "splitter and combiner" at each end of the link is a highly complex filter. As a bidirectional system will be used for both telephony and modulated data; in addition all filtering must be of sufficient standard not to impair telephony transmission beyond the stringent limits given in ITU-T telephony standards, or their equivalent ETSI standards, and handle both the speech signals and the ringing voltage (75v RMS), on/off hook impedance changes and loop disconnect or MF4 dialling. In addition, the telephony system must function when there is no mains power available at the customer's home and so provide emergency telephone access.

Current ADSL technology can reach 5 km or more over standard copper pairs at 2Mbit/s, taking into account worst-case cross-talk effects found within

Downstream in Mbit/s	*Upstream in kbit/s*	*Range in km*
1.5	16	> 5.9
2	16	> 5.5
3	64	> 4.8
6.3	384	≈ 3

Table 2.2: ADSL range

multi-pair cable. For shorter reaches the technology can provide higher bitrates, see *Table 2.2*.

These rates have been actively considered within the ANSI deliberations. Shorter reaches will directly cater for most customers within urban areas in the UK, or for all customers if the actual copper-pair drop is within the range required, and fibre is used from the exchange building to the start of the copper drop. A number of international committees are active in ADSL standardisation such as ANSI, ETSI, and ITU-T. As ADSL could support the Interactive Multimedia Services (IMS) other organisations such as the ADSL Forum, DAVIC and TIA are also involved.

Using ADSL technology over the complete access link is seen as a short term solution to providing advanced broadband services to the residential customer. ADSL allows incremental deployment and is very capital efficient since it can only be provided to those who want it, both attributes are very important in establishing a new field. Once major broadband services have become established (such as VoD) and there is a substantial market for multimedia services in general, then the technical business case for fibre roll-out in the access network can be more readily derived.

ADSL supports both ATM based access and IP based access. Both types of services can be accessed by those with the suitable interfaces. One possible broadband services implementation scenario is to use ATM to deliver MPEG compressed video and IP to support Internet access.

2.3.3 Cable television (CATV) network. Fibre/Coax

The Cable TV companies are currently upgrading their core subsystems (which equates to the first part of the link from local exchange to customers for a telecoms provider) to fibre, leaving the tree and branch coaxial system to serve a small number of homes. Fibre therefore can provide high-speed digital services in addition to any existing TV channels, and these digital services can be broadcast to all the homes served by that fibre, the home equipment picking off the relevant digital information for the service requests.

The cable TV (CATV) distribution system is based on a tree-and-branch topology nowadays and will be based on a star topology in the future to provide higher bandwidth to the end user. The trunk lines are usually made of fibre. High penetration of the CATV is targeted in cities and large communities. Due to the high bandwidth, it has many channels available, which are multiplexed onto the cable using Frequency Division Multiplexing (FDM). CATV providers can use their existing network infrastructure to provide Digital Video Broadcast (DVB) to the end user. Since the number of channels is substantial for broadcast, but not for multicast, interactive services that need a high bandwidth must be considered carefully.

Channel transmission on the cable is primarily unidirectional. Signals are inserted on the downstream channels at the so-called head end. Signals from customer sites are only allowed on certain upstream channels and they are only transmitted towards the head end. Although there is provision for upstream message transmission, many cable systems do not have the actual amplifiers and filters that are needed. In addition, the problems of signal regeneration and noise are harder in the upstream direction as multiple noise sources are merged (noise funnelling).

Cable systems are vulnerable to physical damage, both from ageing and wilful destruction. By ageing all the weak points become potential noise causing spots, creating additional noise ingress and leakage. If the power used in the cable is high in order to limit noise ingress effects, there is more leakage. Furthermore, one big minus for CATV connections is secrecy. If there is no ATM multiplexer in the centre of the CATV star or root of the tree, all data is going on the ATM "bus" which is available to all subscribers. This means more costs, because means to ensure privacy in the connections have to be provided, and ATM does not provide encryption at this level. This has to be done by other means, possibly by providing encryption over the last drop.

2.3.4 Fibre

Currently, fibre is used in trunk lines. A number of cable and telephone companies plan on using feeder fibres to deliver information to the local exchanges, which are serving customers by coaxial or copper pair cables. Traffic in networks is expected to grow quite fast due to new services, and so fibre to the home (FTTH) and fibre to the kerb (FTTC) systems, for example, are expected to be deployed in order to transmit wider bandwidths in the future. At the same time, fibre has other advantages. It has fewer active nodes in the network to maintain and has the ability to be installed in restricted duct space. Furthermore, it is a truly future proof network, but it is now too expensive to invest in fibre in the local loop, on a large scale. The civils costs of installing completely fibre based networks is prohibitive for most telecommunications companies. A large customer base is needed to share equipment costs. Although the new services are anticipated to pay for the upgrade, nothing is certain yet, and most other providers are much more cautious. NTT of Japan is in the process of installing a fully fibre network, not without delays though.

Fibre to the home (FTTH)

A completely fibre based system has been considered by many network providers. BT has conducted experiments on various fibre systems and fibre/copper pair hybrids, in Bishop's Strotford [15], not for VoD services, but for telephony and broadcast service. These were mostly based on Passive Optical Network (PON) technologies. These technologies rely on exploiting the massive potential bandwidth of optical fibre to permit a tree and branch physical cable structure to provide star-type (i.e. independent) services to customers. Although the PON technology used in these trials was not designed to support VoD services, it is clear that with the use of a modified PON technology, or other optical technology, the bitrate requirements for VoD and symmetrical services could easily be provided.

Fibre to the Kerb (FTTK)

In this scenario fibre is installed to the Kerb and not all the way to the customer's home; this reduces civil engineering cost. Form then onwards other technologies such as xDSL will support high speed transmission. In Europe and the USA, trials using fibre to the kerb are being done, and will probably have high speed drops (6+Mbit/s) over a standard copper pair to the customer's home.

Conclusions on fibre

Various options exist, from Cable TV upgrade equipment, with hybrid fibre/copper systems, through to FTTH solutions, none of which are currently optimal to provide full VoD services to a large number of homes primarily due to high installation costs, but all of which have the potential. The precise solution chosen will depend heavily on the capacity and the communications each user will wish to have, in addition to VoD and how the final costs will wash out.

All optical fibre solutions will incur the civil engineering costs associated with laying new cables in the access network, costs which are heavily based on labour charges and not readily susceptible to reduction even in volume. This is why improving on existing technology is a highly desirable intermediate step.

2.3.5 Wireless

Hybrid fibre and wireless distribution networks could be used within a neighbourhood for reducing the installation and maintenance costs, rather than to support mobility. Such a system could be useful in an area where the copper wires are in poor condition or the physical connection between a local distribution point and the customer residents is limited or installation or "rent" cost too much, for example when the service provider owns a fibre infrastructure but not the copper plant to the customer premises.

Wireless transmission of data may be very suitable in some cases, but there are several problems to be solved such as the signal structure in the transition from fibre to wireless and the provision of a wireless reverse channel. Nevertheless, there are some trials for wireless TV distribution networks. Two examples of such an implementation were, first, the British Telecom Millimetre-wave Multichannel Multipoint Video Distribution Service (MMDS) that worked at 29 GHz [16], and second Cellular Visions commercial offering in the New York area under a U.S. FCC pioneer licence in the 27.5 to 29.5 GHz band.

2.4 Network technology

We shall next examine the most important transfer protocols that are appropriate for multimedia transfer. Because the media is contained digitally, the protocol must support digital data transfer. As VoD is a real-time video service, time guarantees must be preserved. ATM, DSM-CC (Digital Storage Media Command and Control) and the future Internet Protocol, IPv6, will all be able to provide Quality-of-Service guarantees.

2.4.1 ISDN (Integrated Services Digital Network)

ISDN is a set of communication transmission standards allowing a single customer line to deliver voice, network services and video in digital form. ISDN was intended to replace POTS (Plain Old Telephone System) and requires specialised central switches, software and other equipment. It provides bandwidth of 64 kbit/s or 2.0 Mbit/s. It is available in most countries of the EU, the US, Australia, Japan and Singapore. Problems have set back its implementation date; and equipment and line rental costs can be high, especially for the home user. Its take-up rate is rather small and has not spread as predicted. In the UK, BT is installing an ISDN line with two 64kbit channels to the residential user with connection charges ranging from approximately £116 to £300 and quarterly line rental of £100-£300 [17]. Demand though is not likely to increase unless the technology becomes affordable for both the business and the home user.

In ISDN, voice and data are carried by bearer channels (B channel) occupying a bandwidth of 64 kbit/s. A D channel handles signalling at 16 kbit/s or 64 kbit/s depending on service type. H channels are provided for user information at higher bit rates.

There are three types of ISDN service: Basic Rate ISDN (BRI), Primary Rate ISDN (PRI), and Broadband ISDN (B-ISDN).

- An **ISDN BRI** line consists of two B-channels (Bearer) that provide 64 kbit/s transmission speed. It also has a 16kbit/s D-channel for control and synchronisation, providing a total of 144 kbit/s data per BRI line. A popular modem operates at 56kbit/s. With the same delivery rate ISDN, being digital (POTS is analogue) can be used to provide a clearer signal than analogue. "Noisy" POTS analogue lines reduce the 56kbit/s maximum modem bandwidth. NISDN-1 is focused primarily on providing basic services immediately to the home user especially.

- An **ISDN PRI** line is intended for users with greater bandwidth capacity requirements. The channel structure is 30 B channels plus one 64kbit/s D channel for a total of 2.048Mbit/s.

- B-ISDN: will support as much as 150 Mbit/s, but will be dependent on a complete optical fibre network. This could be a medium for future high definition television (HDTV) projects.

To access the BRI service, is necessary for the customer to subscribe to an ISDN phone line. Residential customers will also need to use a device called a Network Terminator 1 (NT1). The NT1 performs the multiplexing and converts the 2-wire teleco line to the 4 wire, called S-Bus, ISDN signal.

Many business customers have phone systems that are already digital. In this case, connecting to ISDN lines may or may not require additional hardware, depending on the system.

Many computer workstations are now being shipped with ISDN capabilities. These units connect directly to the NT1, and will integrate voice/data communications through the system software. Most computers though, require a Terminal Adapter (TA). This unit converts ISDN to the serial (RS-232) interface on most computers.

2.4.2 Asynchronous Transfer Mode ATM

ATM is a method for dynamic allocation of bandwidth using a fixed 53 byte packet (cell), known also as "fast packet". The cells use characteristics of both time-division-multiplexing of transmission media, and packet switching of data networks. A "virtual path" (VP) is a unidirectional logical association of VCs, and is set up through the switches involved when two endpoints wish to communicate. A "virtual circuit" (VC), provides for the sequential unidirectional transport of ATM cells, and mimics the circuit switched system. The circuit is established and reserved for the duration of the connection, all packets pass through the reserved nodes. ATM provides a bit-rate independent protocol that can be implemented on several network types and has been proposed by the telecommunications companies as the standard for B-ISDN implementation.

Characteristics of ATM :

1. Scalable technology; potential for extremely high speeds.

2. Cell switching supports both delay-sensitive and conventional data transmissions.

3. Flexible implementation on many media (copper, coax, fibre).

It uses a hierarchical multiplexing scheme where channels can be added or dropped at any point. The coding speeds start from 2.048 Mbit/s to any speed that can be supported by the network.

ATM can be implemented over the Synchronous Digital Hierarchy (SDH) transmission layer. SDH specifies how payload data is framed and transported synchronously across fibre optic transmission links requiring all the links and nodes to have the same synchronised clock for data transmission and recovery. The objective is that products from multiple vendors across geographical and administrative boundaries should be able to plug and play in a standard way. The B-ISDN network will be a truly international network. It standardises transmission around the bit rate of 155.520 Mbit/s which is also called STM-1, and multiplies the bit rate to comprise higher bit rate streams. Thus STM-4 is four times STM-1, STM-16 is 16 times STM-1 and so on.

ATM speeds over a cell-switched network are only limited by the medium and switching speeds, since it is a scalable technology it is not tied to any specific data rate. However, ATM requires wideband fibre/coax cables to exploit its capacity. ATM is planned to be implemented for national back-bone and long-distance carriers and already is being used and tested in many countries. ATM switching technology is evolving to accommodate the special features of multimedia traffic such as bursty traffic, see [18].

ATM and video transport

The network operator is expected to provide a variety of transport types to the video provider. Before sending a video to a set-top-box (STB), the video provider chooses a specific transport type from the network operator. The networking company then establishes a VP from the provider to a STB and delivers the video over the VC with the chosen transport type. The transport types broadly fall into four categories.

1. The first category is Constant Bit Rate (CBR) transport, whereby the ATM network establishes a VC that emulates a CBR circuit between the video server and the viewer's STB. This service can guarantee no or negligible network cell loss, delay and jitter, if the negotiated and obtained bit rate matches the maximum data bit rate required by the application. It is fully characterised by a constant bandwidth, which is negotiated at VC establishment. Note that CBR transport does not require the video server to transmit cells into the VC at a constant cell rate; for good quality transmission the peak transmission rate should be the negotiated bandwidth, if the average rate is used there will be fluctuations in image quality.

2. The second category of transport types is Variable Bit Rate (VBR) transport. VBR transport uses network statistics in order to make more efficient use of network bandwidth. If a VC uses VBR transport, its cells are statistically multiplexed with cells of other VCs in the network links along segments of the VC's route. This transport type can deliver the required bandwidth, but cannot guarantee its availability. Work is being done on how best to implement VBR for video traffic, in terms of rate selection [19], quality control [20] and coder rate control [21]. Compared with CBR transport, VCs with VBR transport are more difficult to manage, both in the network and at the video provider, due to the traffic variability that the scheduling algorithms and the cell admission protocols must accommodate, but can offer more efficient bandwidth management.

   It is important to note that transfer of pre-recorded video is in many ways different from transport of real-time video, such as video conferences or the live broadcast of a sporting event. For real-time video, the exact dynamics of the VBR traffic are unknown to the network. On the other hand, in

VoD the dynamics of the pre-recorded VBR traffic are fully known before the beginning of server transmission and network transport. The network and video provider could exploit this information in order to efficiently manage their resources.

3. The third category of transport types is best-effort. When a VC uses best-effort transport, the network transports its cells to the STB with little if any guarantees on cell delay or jitter. Best-effort transport is only feasible for STB's with large memory and for viewers who are willing to request a start time in the distant future. Best-effort transport however should be relatively inexpensive, particularly if the requested start time is sufficiently in the future to allow for cell transport during off-peak hours, this would require sufficient storage space at the customer end. Intellectual property rights should be considered when provisions are made for storage of material in the customer premises.

4. There are also non-real-time and available bit rate types of transfer, the first is for applications that are not time sensitive and the second for applications that can vary their bandwidth requirements. These types of data transfer provide the operator with more flexibility in network management.

Current proposals for TV and VoD distribution over ATM [22] use MPEG-2 ISO/IEC 13818-1 systems [23] on the codec side over a CBR service class on an ATM network. The traditional CBR service is the simplest solution [24] for video in many applications (e.g. TV broadcasting or VoD) but suffers from disadvantages such as variable video quality (in the case where the negotiated bandwidth is the average not the peak), high end-to-end delay if the average bit rate is negotiated, and relatively high transmission cost [25, 26, 27]. Thus, a more efficient video transport framework will involve VBR video over ATM VBR service class. The VBR service mode allows more and efficient utilisation of bandwidth by using statistical multiplexing [28, 29].

DSM-CC protocol for MPEG over ATM

One of the major factors that will affect the deployment of video-based information services to the home is the availability of open application service protocols [30]. Without such protocols, every service delivered to the home will require its own appliance to receive it. An open protocol allows STBs and other user premises equipment to access multiple service providers. Such an open application service protocol for integrated video-based multimedia is evolving as part of the MPEG-2 standard for encoding/decoding video streams. This protocol called Digital Storage Media Command and Control (DSM-CC) has grown from a few pages in an annex of the recently published MPEG-2 ISO/IEC 13818-1 [31], to over 300 pages in the current ISO/IEC 13818-6 Draft International Standard [32]. DSM-CC provides the protocols needed to deliver a complete application such as VoD, home shopping, or interactive educational video.

DSM-CC began as a protocol for MPEG-2 video stream control (i.e. pause, fast-forward etc.). DSM-CC now encompasses protocols required to support a user being able to find an integrated video service, to download application code, and play stored MPEG-2 video information. DSM-CC contains the use-to-user part that standardises video stream control, directory look-ups, file transfer and database viewing. It also has a user-to-network part. The user-to-network part describes messages for client configuration and downloading and for the transport layer independent initiation of connections over which DSM-CC user-to-user messages and audiovisual content will travel.

2.4.3 IPv6

IPv6 is a new version of the Internet Protocol, designed as a successor to IP version 4. IPv6 was recommended by the IPv6 Area Directors of the Internet Engineering Task Force at the Toronto IETF meeting on July 25, 1994, and documented in RFC 1752, "The Recommendation for the IP Next Generation Protocol" [33]. The recommendation was approved by the Internet Engineering

Steering Group on November 17, 1994 and made a Proposed Standard.

IPv6 is designed to take an evolutionary step from IPv4. It will support directly IPv4 and all applications using it. The limited number of Internet addresses remaining due to its rapid growth, the possible mobility of Internet users and the support of some quality of service, lead the IETF to propose IPv6. The changes from IPv4 to IPv6 fall primarily into the following categories:

- **Expanded Routing and Addressing Capabilities**

  IPv6 increases the IP address size from 32 bits to 128 bits, to support more levels of addressing hierarchy and a much greater number of addressable nodes, and simpler auto-configuration of addresses.

  The scalability of multicast routing is improved by adding a "scope" field to multicast addresses.

- **A new type of address** called an "anycast address" is defined, to identify sets of nodes where a packet sent to an anycast address is delivered to one of the nodes. The use of anycast addresses in the IPv6 source route allows nodes to control the path over which their traffic flows.

- **Header Format Simplification**

  Some IPv4 header fields have been dropped or made optional, to reduce the common-case processing cost of packet handling and to keep the bandwidth cost of the IPv6 header as low as possible despite the increased size of the addresses. Even though the IPv6 addresses are four times longer than the IPv4 addresses, the IPv6 header is only twice the size of the IPv4 header and extension headers of unlimited length can be added if required.

- **Improved Support for Options**

  Changes in the way IP header options are encoded allow for more efficient forwarding, less stringent limits on the length of options and greater flexibility for introducing new options in the future.

- **Quality-of-Service Capabilities**

  A new capability is added to enable the labelling of packets belonging to particular traffic "flows" for which the sender requests special handling, such as non-default quality of service or "real- time" service.

- **Authentication and Privacy Capabilities**

  IPv6 includes the definition of extensions which provide support for authentication, data integrity, and confidentiality. This is included as a basic element of IPv6 and will be included in all implementations.

The most important feature for the IMS services in the IPv6 are the Quality-of-Service Capabilities which make the delivery of real-time applications possible.

Special fields will exist in the IPv6 header, the Flow Label and the Priority fields, which will be used by a host to identify those packets for which it requests special handling by IPv6 routers. This capability is important in order to support applications which require some degree of consistent throughput, delay, and/or jitter, commonly known as multimedia services.

IPv6 implementations are being developed for many different host operating systems and routers. This includes host implementations by Apple, Bull, Dassault, Digital, Epilogue, FTP Software, IBM, INRIA, Linux, Novell, Mentat, NRL, NTHU, Pacific Softworks, Process Software, SICS, SCO, Siemens Nixdorf, Silicon Graphics, Sun, UNH, and WIDE, and router implementations by 3Com, Bay Networks and Cisco Systems, Digital, Hitachi, Ltd., Ipsilon Networks, Merit (routing protocols), NTHU, Sumitomo Electric, and Telebit Communications. More information on IPv6 is available from [34].

2.4.4 ATM and IP

ATM and IP are the most prominent examples of the convergence of the telecommunications and computer industries [35]. ATM was originally conceived in the telecommunications field and has been developed under the influence of the computer networking style in the ATM Forum. IP was developed in the computer world to support the shift from data-only to multimedia information transfer. IP has been one of the key components in enabling the growth of the Internet and its popularity is still increasing. Introducing QoS guarantees adds to its usefulness and versatility [36] towards an Integrated services network [37]. IP adopts a connectionless approach, while ATM is connection oriented. IPv6 will be able to support end-to-end delay guarantees, whereas ATM is capable of doing it now.

On the one side, the telecommunications community is embracing the Internet protocols as both the foundation of computer communications over carrier networks (IP over ATM) and the vehicle for client-server interactions in novel telecommunications services, such as VoD. On the other hand, although the Internet community is sponsoring ATM as the promising high-speed bearer technology of subnetworks, the vision of the Internet is still focusing on the protocol suite to be executed exclusively among hosts at network edges and routers at subnetwork boundaries. Their similarities have lead to a number of proposals for integrating ATM and IP. This integration will establish a way of keeping the advantages of both protocols and support further the implementations of multimedia services.

2.5 Summary

In this chapter we have presented the most significant of the support technologies for IMS, and specifically VoD. MPEG and JPEG image compression offers substantial savings in transmission bandwidth. It covers a range of video qualities from low (VHS equivalent) to high (HDTV) according to the user demands and network support. MPEG decoders can be purchased in the market from a number of manufactures who are active in the field. Standardisation of the latest MPEG versions -4 and -7 is still active. MPEG encoders are more expensive, as encoding is an intensive computing process.

The technology employed currently in the physical network layer can be used for the IMS services thanks to ADSL and its versions. ADSL modem costs are still high (£1000) but prices will drop with demand. Fibre represents the only future proof solution but its installation cost is prohibitive at the current stage.

ATM and IPv6 are presented as the two strongest candidates for transporting multimedia data to the end user. ATM was developed as a transmission system for ISDN traffic. ISDN has started to gain market share, but its pricing still makes it unattractive for the residential user, especially in the UK. In Germany and the US for example prices have dropped considerably and hopefully this trend will be followed in this country.

In the next chapter we present a number of VoD field trials done by the leading network and entertainment providers. These trials were more market research oriented. Their main target was to evaluate the potential customer demand for interactive services. Although limited on customer base and offered material, they provided some very useful results as to the form of an interactive service.

Chapter 3

VoD Trials

3.1 Introduction

In this chapter a number of VoD service trials are presented. These were based on specifically constructed equipment and were conducted on a restricted customer base ranging from testing company employees to a few thousand paying customers.

Finally we present our view on system design for IMS (Interactive Multimedia Services) using VoD (Video-on-Demand) as an example.

3.1.1 The legal issue

Telecommunications companies were, until recently, state owned and operated industries in most countries. In the post-privatisation era, laws regulating their operation forbid them from broadcasting using their network. When privatisation schemes started in the early 80s laws regulating the transition, excluded telecoms companies from entering the multimedia services field.

In Europe, after negotiations with the telecommunications companies, legislation is changing. The EU has published a major consultation paper on regulating a telecommunications world in which the boundaries between IT, telecoms and broadcasting are fast breaking down. The convergence of telecommunications and IT is evolving rapidly and new laws are needed for the emerging environment. Telecommunications operators will soon be able to use their network for other services even for multicast services. With complete market liberalisation, in all EU countries by the end on the century, Cable TV operators and telecommunications providers will compete for the same market, this is hoped to provide more choice to the customer and bring prices to reasonable levels.

In the UK, the first country in the EU to privatise its telecommunications company under The Telecommunications Act of 1994, the regulating authority, OFTEL, is contributing to the development of policy in the new framework of competition. OFTEL regulation is aimed at preventing abuse of market power and anti-competitive agreements. The provision of conditional access services is to be done on a fair non-discriminatory basis. OFTEL as the regulator is responsible for enforcing these principles.

The rest of Europe is somewhat behind as it lacks OFTEL's experience in the field. As of the 1st of January 1998 the European market has been liberalised. In some countries though the national providers try to keep their share by exploiting their monopoly status, for example in Germany the national provider Deutsche Telekom was asking for a change-of-provider fee of $ 53 [38] in the beginning of 1998.

In the US, the congress passed the Telecommunications Act of 1996 on February the 8th, by a large majority. Under the new law, services are allowed or disallowed not by regulation, as in the past, but by contracts negotiated among the providers, none of which is a regulated monopoly. The effects of these revolutionary changes will be felt for years to come. Cable television companies may offer telephony and information services, and the telephone companies can carry information and entertainment services into homes. Lastly, media companies have greater freedom to acquire and hold assets in any one service area.

The telecoms market for 1996 was estimated to be 400 billion pounds. The possible introduction of new services will boost its value even more. All major telecoms companies in the world have been experimenting with interactive multimedia service provision. The idea though is not new, from the 50s companies such as BT have tried to build systems that can provide on line information. In the UK, BT tried to provide an information service called Prestel (a form of and forerunner to teletext) using its network in 1979. The service did not become very popular for a number of reasons. The interface was difficult and slow to operate, the cost of the required equipment was left to the customer and the information content was rather "dull". Although BT's system did not become very popular, in France, France Telecom managed to roll-out a service which reached a large customer base of 6 million homes and businesses, and could be described as the forerunner of their on-line interactive services, the Minitel. Minitel offers on-line services, such as directory enquires, teleshopping, games, ticket reservation etc. The system was to be financed by cutting the cost of printing paper catalogues which it was to replace. The Minitel was more successful than similar services launched in the UK, Germany (the service was called Bildschirmtext), Sweden etc. The modem and display device was provided and a friendlier interface made browsing easy. It became an indispensable service for those using it. The Minitel gave the first insights as to what should be considered when providing interactive broadband services.

With the possibility of video delivery, the market potential of the new services becomes immense. This is why not only telecoms companies, but CATV companies and even entertainment producers are testing their potential entry to the interactive multimedia market. New alliances are formed which aim to survive and dominate in the free market of the future.

Most of the field trials to date were primarily market research rather than technology trials. The take up rate of the new services, the user's willingness to pay for what as well as how much and the least required functionality, have been the main topics of research. Next, a few of the most significant of these trials are presented, along with a discussion on their findings. Information on VoD trials is not widely available due to the commercial aspects of the topic.

Technology issues seem to have delayed and downsized many of these trials.

3.1.2 BT

British Telecom (BT) is very interested in VoD. Because it was not allowed to offer broadcast type entertainment services using its network, BT has conducted a number of trials focused on the possibility of providing interactive broadband services. Currently BT is being motivated to provide Internet access to schools in exchange for the right to broadcast using its network [39].

BT was heavily involved in cable-TV systems and technology in the early/mid 1980s and installed, for the period, a highly advanced cable-TV system in Westminster based on a switched star technology. The switched star network (SSN) architecture made possible the switching on-demand of a video channel from the head end of the system to a specific customer from a pool of available channel capacity. The first major exploitation of this capability was intended to be for an on-demand one-to-one video library service. The video library system was developed, and a small-scale system installed on the Westminster cable system. It used analogue technology for storage, transmission and switching of the audiovisual signals.

A similar small-scale on-demand video library was also installed for the Bishop Stortford fibre-to-the-home trials.

BT has invested heavily in Interactive TV trials in recent years. This is not only to progress the technology, but more importantly to start to understand how customers will react to such services. They also aim to understand and build relationships with some of the other key players involved in an Interactive TV system, namely the Content and Information providers.

The 1994 technology trial.

A small-scale field trial in 1994 of basic VoD services was carried out to 60 employee customers in their own homes, in Kesgrave near Ipswich, Suffolk. The service offered was fully on-demand one-to-one viewing of old TV programmes and a few films. The objective of the trial included testing the end-to-end system technology within an operational environment.

The set-top box was based on standard PC technology with the keyboard, screen and mouse removed, and dedicated video compression hardware added.

The server held nearly 70 hours of programmes on-line, stored on RAID-based disks. Material was changed during the trial in order to keep the service attractive and also to gain experience of the problems involved in the process.

Interactive TV 1995 market trial

The aim of the Interactive TV market trial was unashamedly to find out what customers want and how much they were willing to pay for various services. The trial took place in the Ipswich/Colchester areas in East Anglia. The trial involved 2,500 paying customers. The customers were carefully chosen so as to be a representative demographic sample of the UK population.

Customers were connected over ADSL or passive optical networks. One server provided service to all customers, and transmission from this server to the local exchange buildings was ATM over SDH.

The services were ordered using the standard remote control unit and delivered to the user's existing television over copper pair or fibre. The STB was an Apple computer, specially adapted.

The service involved more than 100 service and content providers, including major Hollywood and European movie studios [40].

Westminster video jukebox trial

In parallel with the above Interactive TV trial, BT was also investing in a trial of movies and TV programmes on demand over its existing Westminster cable-TV system. This trial was called video jukebox and tested customers' reaction to on-demand programmes within a cable-TV environment. The trial started in late 1995, growing during 1996, to a minimum capability of 200 titles on-line. Technically, the new system being installed replaced and upgraded the old

Video Library system from the mid-1980s.

3.1.3 Bell Atlantic

Bell Atlantic has been involved in many trials, a recent one (March '95) was conducted in Fairfax County, Virginia. This market trial was the second phase of an ongoing product development effort that begun with a first phase, technical trial in 1993. The product was designed to be commercially viable (i.e. key components are cost effective), but at the time was not viewed by Bell Atlantic as being market ready. Market readiness required further product development and refinement, in key areas such as the number of users it could support and the movie title range it could offer.

Product Description

It included 655+ archived programme choices with 25 percent monthly refresh rate. Customers paid for each programme watched, with prices ranging from $ 0.49 to $ 4.49. The service was targeted as a substitute for pay-per-view and video rental.

Four programming categories were offered: entertainment (movies, TV shows); kid's zone (movies, TV shows, educational programmes); learning and lifestyles (documentaries, how-to's); home shopping, as well as, a spotlight category (promotions, programming packages). Additionally, users were offered full VCR functionality. They could, stop, fast-forward and rewind the movies at will.

In the transmission area, ADSL technology over existing copper phone lines was used. The ADSL equipment was manufactured by Westell and allowed simultaneous use of telephone and video over the same line.

The major findings of the trial were that users would welcome an alternative option to the video rental store and even were willing to tolerate slightly higher prices. The major problems were due to disfunctioning equipment in the customer premises. New release movies attracted 50% of the demand on the peak time from 8 to 11 in the evening.

3.1.4 Conclusions from VoD trials

Up-to-date, all VoD trials conducted have used a very limited customer base. A few hundreds to a few thousands is the usual number of users quoted. These numbers are quite small compared to the expected take-up rate of multimedia services. The expansion to larger geographical areas needs to be taken into account as it could provide new challenges. The server manufactures were usually nCube and the disk provider Oracle, although other providers were also used. There were cases were actual dedicated VCRs were used. The STBs were specially adapted PCs from a number of providers such as Apple, IBM, Zenith, Philips and others. Most trials used ADSL/ATM networks or Cable networks.

The most important conclusions can be summarised as follows. With respect to successful roll-out of the service, the STB needs to be low cost for the end user. It can be subsidised by the provider, as was done with Minitel and cellular mobile handsets. As mass production decreases prices, this appears to be a possible option for the service provider.

The price range of the material offered should be similar to the prices offered by competitors, such as video stores. The price of a two-hour phone-call can be substantially higher than the price of renting a movie. The long connection times should be considered by the network operators, when upgrading their networks. There is willingness to pay for network delivery, from the customer's side, but this does not total the current price of a two-hour phone call.

Consortia are forming including network, content and hardware providers, in their effort to become the key players of the future. A few of the companies that have conducted such trials are: BT, Deutsche Telekom, France Telecom, Swiss Telecom, Telecom Italia, Svenska Kabel TV, Ameritech, Bell Atlantic, Bell South, Pacific Bell, SouthWest Bell, Sprint, TCI, Time Warner, Telecom Australia, Hong Kong Telecom, Israel Telecom, Korean Telecom, Singapore Telecom, NTT, JSAT.

3.2 VoD Service Elements

The most important VoD service elements have been identified to be the following:

1. **The server.** The server represents the storage and playing device. It can be one integrated device or a combination of video playing and storage machines.

2. **The transfer network.** This represents the connection from the server to the user. It includes the core network as well as the last-drop to the user premises.

3. **The Customer Premises Equipment** or what is more widely known as the Set-Top-Box. The STB is all the equipment needed at the customer end in order to access the service. Its cost could be subsidised by the service provider.

3.2.1 The server

We have two types of server: the storage and the playing or cache server. The storage server is the device used for storing the offered material. The material is likely to be stored in its compressed form. There are four media of storage, differing in cost and access speed. They can be used for different classes of data according to its popularity length etc. We have the fast access medium such as RAM, which can be used for frequently accessed material and yield fast access times. We have optical disks, such as CD-ROMs which can be used for frequently accessed material and slightly slower access speeds. Disk arrays have also been proposed such as RAIDs, which can give higher fault tolerance to the system. Then there is tape which is very cheap but has slow access speeds and does not allow parallel access and search functions. Tape can be used to archive seldom accessed material. The storage server should be able to provide the service network with multiple access streams. This will facilitate the fast downloading of material to the cache servers. The higher the output rate the more efficient the server but the more expensive it will be. This should be determined by the service needs and the construction costs.

A centralised or distributed solution is possible. In the centralised solution all material is stored and accessed from one central location. A high transfer rate will be required to enable management functions in parallel with movie traffic and make the network more fault tolerant. In the distributed solution, information may be stored nearer to its access site. This will reduce backbone network bandwidth take-up. Extra storage capacity would be required for the mirrored material. Mirroring and distribution into the storage servers could follow popularity trends.

The cache server is the server accessed by the user. The material after being downloaded from the storage server can be accessed by the service users. VCR functionality and a substantial number of simultaneous users are the most important features the server should support. Technologically it is the most difficult part to construct. Real time interactive access imposes the most demanding terms. Processing and access speeds should be sufficient to allow for the service to be called on demand. Additionally, end-to-end delays due to the transfer network may be added. All major parallel processor companies have been involved in constructing VoD servers, such as nCube, IBM, Digital etc. But, to date their capacity could not satisfy a broad customer base and they are quite expensive. Clustering is considered as a potential solution for increased capacity. Prices are expected to fall when the application is widely rolled-out.

3.2.2 The transfer network.

The transfer network is divided in two major categories. The core or back-bone network and the local loop network. The core network should be able to support the vast volume of traffic generated by the new services. Optical fibre is already is use. Most of the back-bone traffic in Europe and the US is routed through fibre. ATM and IPv6 are likely to be the key technology for back-bone networks.

Satellite is also coming into play, especially for Internet as Intel and Microsoft have announced investment in satellite networks for providing Internet access.

The local loop network is the last mile of wire over which information has to travel in order to reach the customer. This last mile has been custom constructed for one thing–voice telephony–and its structure has never changed over the past 50 years. Now with the introduction of digital technology and the dawn of the information age, the possible bottleneck of the local loop network has been brought forward. The answer of technology is twofold. Efficient compression standards exist and x-DSL can deliver the compressed data to the users. Limitations exist on the distance over which x-DSL can operate and the bandwidth it can support. It offers though the considerable advantage of low cost installation making use of the existing infrastructure.

3.2.3 The customer premises equipment (CPE)

The CPE or STB, see *Figure 3.1*, is the term used for the equipment required by the user in order to receive the interactive multimedia services. There are two directions for designing the CPE. One is to have a highly intelligent piece of equipment, which could probably store data and perform other functions. This will be rather expensive and have the capabilities of a top range PC. The other trend is for the STB to be as simple as possible, and therefore as cheap as possible, and easily affordable by most users. This will impose higher requirements on the supporting network. The third solution will be an integration of the PC and the STB into one device with possibly two monitors attached. One could be the TV set. All solutions should support real-time MPEG decoding. A cheap STB will help increase the take up rate of the new service. A more expensive STB will provide more features and probably better quality, at least in the initial stages of the roll out. An example of a cheap STB, which was provided for a small rent by the service provider, is the STB for Minitel in France. France Telecom was the only company to provide it freely to the customers who wanted to join the service, this is one of the reasons the Minitel survived and succeeded, when other such experiments in the UK and Germany failed.

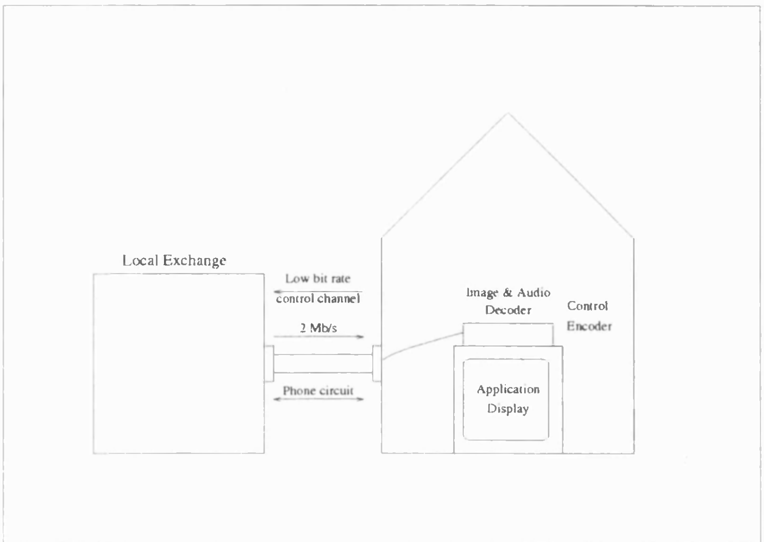

Figure 3.1: Customer Premises Equipment

The STB, interfaces the home TV into the video subscriber network, decompresses video and converts it into standard TV format. A VoD STB will provide VCR-like functionality, allowing the user to rewind, fast-forward, slow motion play etc. by sending user commands upstream via the channel. Although memory is quite cheap, it shouldn't be used excessively as its cost will contribute to the total cost of the CPE. The CPE should be priced competitively with respect to its existing rivals: CATV & DBS (Direct Satellite Broadcast) services and VCRs.

3.3 Summary

In this chapter we presented a number of VoD trials with the limited data that are available from the organisers. VoD has attracted the interest of all possible participants of Interactive Multimedia Services (IMS), the network operators (telecommunications or CaTV), the content providers (i.e. Disney) and the hardware manufacturers (nCube, Oracle etc.). Most trials served as market studies on customer demand and expectations and were conducted with a small customer base.

In the next chapter we present our design for a VoD server, in a distributed network environment. The server presented is a playing server, which provides the end user with a high degree of interactivity, depending on the service implementation. It can offer service to a broad customer base and a network of these servers can be used to upgrade the existing telecoms or CATV network infrastructure to a fully interactive multimedia network.

Chapter 4

VoD and Multimedia Server Design

4.1 Introduction

Following the discussion of system technology in the previous chapter we now present our design for a VoD server in a distributed networked implementation. After a brief discussion of the advantages and disadvantages of the distributed solution, we continue to present our VoD server design. We start by presenting the most important characteristics of the cache server, in terms of support of the user access pattern and interactive functions capability.

We present our idea of segmenting the material offered from the server and loading it onto independently accessed memory units. The user connection is transferred from segment to segment as the user watches the whole movie, but can also have any interactive function since the connection can be transferred to any segment at will. As we require a powerful server, which should be able to support a substantial customer base, we go from a simple space switch, to a multistage switch. The multistage switch can be less complex, simpler construction though, may introduce a blocking probability on server access. This blocking probability can be kept at acceptable levels with appropriate switch design. To reduce the control complexity at the segment access level we create a two stage control mechanism.

Next we examine the digital implementation of our server and how a TST (Time-Space-Time) switch would work. We then analyse the way the interactive functions would be provided.

4.2 Centralised versus distributed implementation

The VoD service cannot be clearly categorised as a centralised or distributed service. It has attributes that are served best by aspects of both a centralised and a distributed topology. The high demand movies should be distributed to facilitate easy access and reduce core network bandwidth take-up; the less popular movies can be easily offered on a centralised basis. Storage can be either distributed or centralised too. Server resource usage is thus kept to a minimum, and core network bandwidth usage is not high. The discussion over centralised and distributed solutions has been long and intense. Both solutions have their advantages and disadvantages, some of which we shall present here. But in the real world nothing is purely black or white, therefore a solution using both approaches, as appropriate, is optimum.

As already outlined one of the most important service elements is the multimedia server. In the concept of the multimedia server, two possibly distinct servers are discussed: the storage and the cache server.

4.3 Cache Server

The cache server design is the most demanding in terms of processing power and user access speeds. Here we are focusing on the design of the cache server which shall be referred to as the VoD server.

The exact requirements of a VoD server are not yet fully known. The field of multimedia interactive services is very new and no traffic and usage models have been established yet. By using what is already known from conducted field trials [41], video store statistics [42] and "logical assumptions" on service characteristics, the basic requirements can be summarised as follows:

- **Independent access.** The user should be free to use the facility without suffering from the actions of fellow users.

- **Accommodate a range of access patterns.** The very popular movies would require a massive one-to-many connection, which should be presented to the user as independent and non-blocking. Some 40% of the customers will be watching the same movie during peak time [41]. As the system is destined to support a large customer base, the number of users served per server will be crucial in determining total service capacity. Copyright [43] should also be taken into account. Flexibility of movie delivery and use should allow for the protection of the interests of the intelectual property holders. At the other extreme we have the rarely accessed titles, which would require a one-to-one connection with interactive functionality and can in principle consume the same server capacity as the popular ones, without rendering the same profit to the service provider.

  These requirements impose further specifications on server design. How well the service performs with respect to the above requirements depends on the service design and on the actual service contract negotiated between user and provider. The efficient and effective balance between varied content availability and profitable running of the service will be the key factor in rolling out and establishing a successful VoD and multimedia service. As already stated, the server is one of the most important parts of the service. Its functionality will be directly linked to the functionality of the service delivered to the end user. Different capacity servers for different classes of material are a plausible option. Their interconnection and communication will become an integral part of service design and essential for the effective running of the service.

- **Interactive capabilities available to the end user.** In order to provide the end customer with true VoD, interactivity should be perceived as "instant". The delays due to server response time will be added to the delays due to the transport network. The total of the two delays should be kept to a minimum. If this process is not successful then the service will be severely impaired and reduced to near-VoD. The support of interactive viewing, will in turn be mirrored in the set-top design requirements. Reducing server delay as much as possible, allows for compensation of delay fluctuations which might be incurred due to the transport system, in a packet switched network, which are hard to control due to the large number of factors affecting them, such as traffic statistics and potential equipment fault. If the whole connection from user to service provider is circuit switched no such problems are encountered, but if part of the connection is packet switched, then the service must be designed carefully to avoid unnecessary delays from this part of the connection.

4.4 The VoD server design

Here we present our design for a multimedia and VoD server in a distributed service environment. In our proposed design the cache server is placed close to the end user, the storage server may be centralised, supported by a network of storage servers for the most popular material. The servers are placed at some point in the communications distribution network for example, at the local exchange level.

Four classes of VoD service have been identified, depending on the degree of interactivity provided to the end user.

- Pay-per-view (PPV), in which the user signs up and pays for a specific programme, similar to the existing CATV PPV services. The programme is delivered without interactive control functions.

- Quasi VoD, in which the users are grouped based in a threshold of interest; if demand does not exceed the threshold, service is not delivered. The material is multicast in relatively large time intervals, such as 15-30 minutes; users can perform rudimentary temporal control activities by switching to a different group.

- Near VoD, in which interactive functions such as fast-forward and rewind are simulated by transitions in discrete time intervals (in the order of a few minutes). This capability can be provided by multicast of the movie in fixed time intervals of a few minutes.

- True VoD, in which the user has complete and interactive control over the session presentation. The user has full VCR capabilities, including pause, forward-reverse play, freeze and random access. True VoD is a true one-to-one service.

In this work VoD is used for true VoD based on human perception as to the threshold above which delay impairs real time interactivity; as a superset of all other types of VoD. The ideas can be used on the other types of VoD as well.

4.4.1 Data Striping

The VoD server should support multiple accesses and interactive actions. Each user should be able to have an autonomous and independent connection to the server. Having one dedicated copy for each user would reduce the service to the level of remotely placed video players, which would be uneconomical for the service provider. Having one copy for all users is broadcast and not dedicated interactive service. To overcome this problem we have introduced a variant of the idea of using one copy of the movie for all users. Here the copy is not used as a complete entity, but is divided into small movie segments. Each segment is a small autonomous unit. This unit can be accessed by any number of users the server can support. The user connection can move through these units at any order any rate. The degree of interactivity depends on segment size. Thus, we manage to provide interactive and independent service using only one dedicated copy of the movie. This is similar to the idea of striping data across disks in order to emulate higher speed disks. Data striping has been used in other work as well [44, 45, 46, 47]. Here we present a generalised approach based on this technique, which provides full VCR functionality. In our design, the movie is divided in segments and each segment is loaded on a separate playing device. The segments are accessed by a number of concurrent users. Although the users have concurrent access to segments, they can using appropriate comtrol scan the segments at any rate other than the playback rate and have the possibility to access any"allowed" point in the segment. The segment is essentially multicast. Striping the movie across small segments provides the user with interactive functionality.

This element is unique in our design. We believe the functionality it provides makes it essential for a true VoD service implementation. Depending on segment size this technique can be used for a range of VoD services, from N-VoD when segment size is in the order of minutes, to true VoD when segment size is in the order of seconds retaining its advantages of access and availability, throughout the range.

In the servers, information received from the storage servers is buffered and then loaded onto server memory inputs, the segments. This material could be, for example, magnetic disks. Movie segmentation allows a number of users to access it simultaneously, as each user is effectively receiving a multicast of the relevant segment.

The segments are played simultaneously and in synchronisation. This ensures that as one segment finishes the next one is starting. The size of the segments is a compromise between server cost and the maximum interaction delay allowed. As will be later explained, the average interaction delay will be half the segment size. This delay is directly related to the QoS supported by the connection. The shorter the staggered broadcast interval the higher the quality of the offered VCR functions especially, the pause function. The service provider should balance QoS offered to the end user against server cost, since good QoS requires short segments. In *figure 4.1* the decomposition of the video into segments is represented as separate databases. This is not essential but is conceptually easy to visualise.

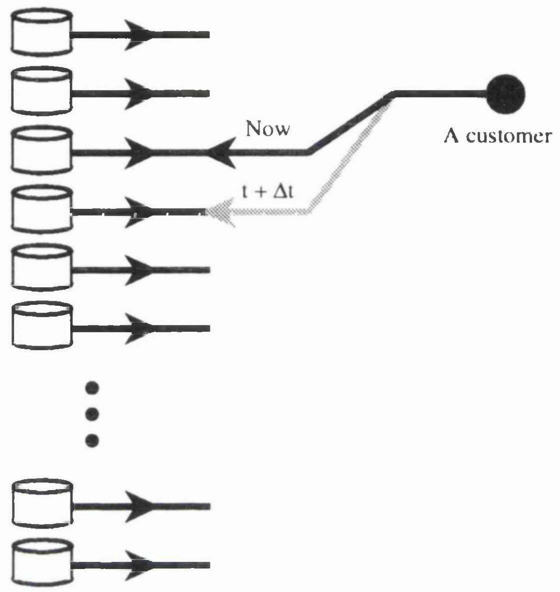

Figure 4.1: Video Switch with movie split into segments

A space switch is connecting user channels to segments. As the customer watches the movie, their connection is being switched from segment to segment.

The memory inputs can be loaded with the compressed or normal version of the movie. This depends on the speed and cost of the encoding process. The VCR functions are affected too. Currently, viewing a fast-forward version of a movie in MPEG is not of very high quality; in reality, the user receives a number of separate frames. This though, is often acceptable. Loading the compressed version of a movie, will make loading an even faster than real-time process. Storing the movie appropriately (i.e. parallel disks), so that we can load the segments in parallel can reduce the system response time.

The video delivery system must have an interface onto the communications network. This is done through a basic switching process that is provided within the current network. Each segment can be viewed simultaneously by as many customers as the switch can process simultaneously, *as shown in Figure 4.2.*

It is multicast in essence at this point and requires no interaction with the customer. The only requirement is a signalling interface into the switching system, which controls the segments to be viewed. Our design is based on existing equipment to which new control software is added. This signalling interface provides the customer access channels with the correct movie frames, thus preventing multicast of frames into the network. The result of this is that the switching interface can connect the customer inputs onto any of the segments and there is no contention, provided the number of connections does not exceed switch capacity. The user views the movie as if he had a completely independent connection to a dedicated copy of the movie. Additionally, the switching system can handle each connection independently and move it through the segments in any order. This functionality gives the user the ability to randomly access any part of the movie which constitutes the most basic degree of functionality available to the end user.

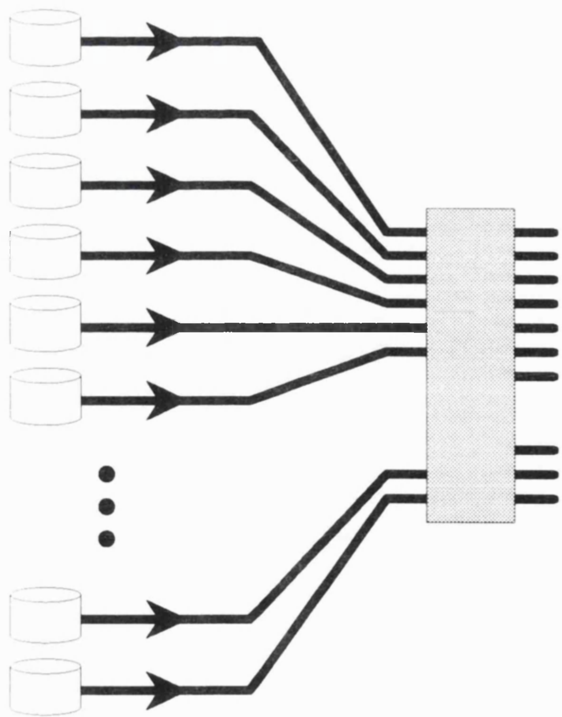

Figure 4.2: To supply a large base of customers a switching interface must be provided onto the delivery system

This switching process is shown here to be a space switch. In *Figure 4.3* the VoD switch is depicted as an ordinary space switch. Again this is not binding but simplifies the explanation. Next a brief discussion on how a space switch operates is presented.

A space switch is used to switch the customer lines or channels to the correct switch inputs. In the VoD server switch, the customer lines are the customer channels and the switch inputs are the movie segments. At some arbitrary time a customer wants to access the VoD service. As soon as demand for the service arises the customer is assigned one of the customer lines. To watch the movie interactively the customer needs to receive the correct segments of the movie in the correct order, and also be able to pause, fast-forward *etc.* at will. For the switch operation, this requirement is translated into switching the customer line from segment to segment sequentially and additionally accessing a certain part of the movie if instructed to.

In the distributed model approach, after having distributed the interface connections, then the complete service networked structure can be considered. For instance, a small number of storage delivery systems could be used, whilst placing a large number of interfaces in the network at the local exchange. This means that the video delivery is local as far as the communications network is concerned and it puts no strain on the bandwidth usage in the core, except for movie downloading which is not done very often. Such a distributed server system will require a management scheme to load movies and establish connections. We will present such a scheme in the next chapter.

4.4.2 The provision of VCR functionality using the proposed design

The compressed movie is loaded as video segments. An MPEG compressed sequence of frames consists of I, P and B frames. The I frames are essentially still images compressed and used as reference points for the P and B frames. Their decompression does not depend on previous or following frames. A possible implementation should co-ordinate the number of I and P and B frames, so that the beginning of a segment is always an I frame. This will provide the user with random access.

- **Fast-forward-rewind.** These functions may be facilitated by the STB cache memory. Storing future or past frames in the STB cache memory, will facilitate fast-forward-rewind actions. The number of these frames should be decided by the provider and compensate for switching and network delay. If the user chooses one of these functions, these frames will be used immediately. Care should be taken if the duration of the function exceeds the stored frames, the new frames will come directly from the server, which should change the rate the connection is moved through the segments. Alternatively, a fast-forward version of the movie, may exist alongside the normal one. As the user requests one of these functions, he is connected to the fast-forward version, until he resumes normal viewing again. Small sized segments may also be used.

- **Pause.** Pause is the most essential and least demanding function to realise. When pause is selected, the user connection is temporarily stopped. The established circuit maybe preserved or released, depending on the pause state duration. If the average expected pause state duration is exceeded the circuit may be released. The current movie segment is noted. When the user wishes to resume movie watching he is connected to the segment he was disconnected from when he issued the pause command. When movie watching is resumed, a few frames will be sent again to the end user. If the size of the segments is relatively small, in the order of a few seconds, the interaction is perceived as "instantaneous". If the segment size exceeds this limit, the user will unavoidably experience a delay, or review a number of frames, until he may continue normal viewing. For material where instantaneous functionality is needed, a small segment size is essential.

- **Random access.** The connection is moved an arbitrary number of segments forward or backward, the user receives the I starting frame of the segments.

Due to the segmentation technique, there will a delay when starting the viewing process. This will happen in both cases when accessing the movie and when resuming form pause. The average delay D_{av} will be half the segment duration.

For less popular movies, where reduced functionality is acceptable, segments of longer duration can be used. For a highly interactive service the perceivable delay duration used may be within the limits tolerated by viewers, in the order of seconds.

4.4.3 Space switch limitations

Although the space switch approach offers random access and video player functionality, it does not support a large customer base. A huge spatial switch would be needed. The complexity of the space switch C in terms of the number of cross points, measured as the product of inputs M times the outputs N, $C = M \times N$, increases considerably as new inputs or outputs or both, are added. As a consequence, it becomes increasingly harder to construct a larger space switch.

The major disadvantages of this design are space switch size and control complexity. The realisation of large space switches becomes quite uneconomical. Furthermore, the control and processing required for such a structure increase considerably, to an extend that their realisation would be almost impossible.

A switching process should be used to move the customer connection from segment to segment, responding to the control messages from the user. A space switch offers a limited number of cross-connections, which may not be enough for the proposed application of VoD. Additionally, it offers very poor accessibility. Connections are possible from slot k in multiplexed input, only to slot k in the multiplexed output, *see Figure 4.4*, thus no data rearranging is possible. As a

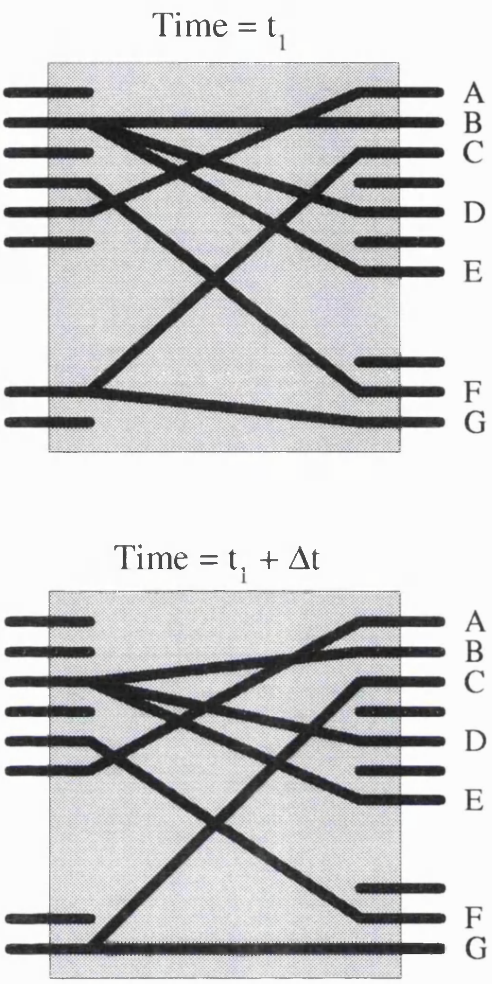

Figure 4.3: Video Switch Block

result, a space switch is not usually employed on its own, but is used with a time slot interchange switch [48], such a switching process we propose for VoD.

4.4.4 Multistage Switch

A space switch connecting M inputs to N outputs has complexity $M \times N$. This means that as we add inputs and outputs to the switch, its complexity increases even further, making its construction quite expensive. If we want to simplify the component structure and keep the blocking probability very low, the common approach is to build a multistage switch. This is in effect a network of switches.

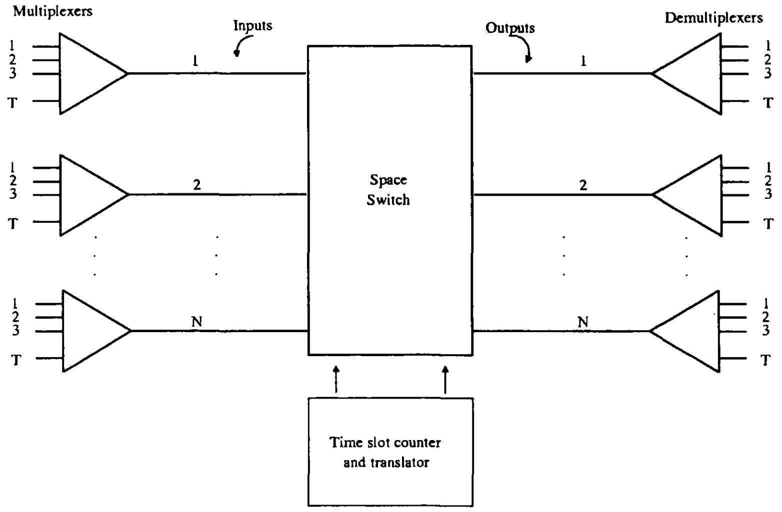

Figure 4.4: A space array used to provide space switching of digital signals

A multistage switch is used to provide a less complex version of a single stage space switch *Figure 4.5.* In the multistage switch more than one layers of space switches are interleaved between the input and output of the switch. Constant circuit paths are constructed between the switch input and the switch output. The input and output of the switch can also be divided into more than two switches.

Multistage switches may be blocking depending on their design. Blocking probability is the probability that a link between any switch input line and any switch output line will not be established at a given time due to some part of the link being used by another connection. The less the blocking probability required the larger and therefore more complex the switch. By setting accordingly the number of intermediate stages and the size of the switches we can control the blocking probability. Most modern switches have a very low blocking probability and are called *essentially non-blocking.* Here we must state that the switch intended for a VoD application should perform at least as *essentially non-blocking,* especially to the already connected users, since blocking during movie watching will amount to bad Quality-of-Service (QoS). A VoD switch should not terminate existing connections due to internal path rearrangement and guarantee the normal termination of the existing connections. If abrupt connection termination occurs during connection rearrangement this will not be tolerable by the end user.

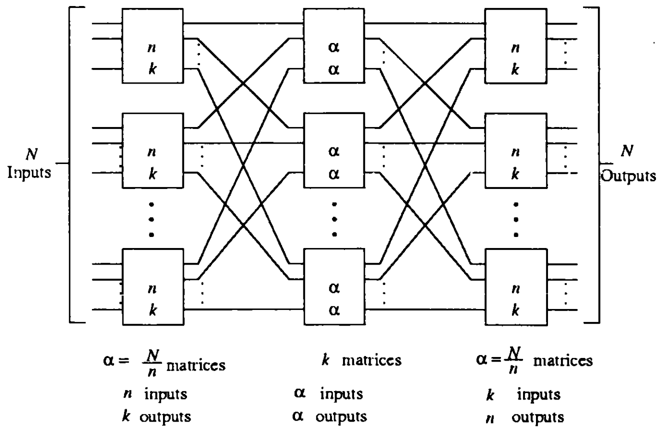

Figure 4.5: Schematic representation of a three stage switch.

If we have N input lines at the one end and N output lines on the other end (a symmetrical switch), we can divide the N input lines into N/n groups of n lines each. Each group of n inputs is accommodated by an n-input, k output matrix. The output matrices are identical to the input matrices except they are reversed. The intermediate stages are k in number and have $\alpha = N/n$ inputs and α outputs. It is observed in Figure 4.5 that an arbitrary input can find k alternative paths to an arbitrary output, because the connection can be established, through any one of the k intermediate stages.

To determine the number of cross points or complexity, C, in a three-stage switch we note first that each of the first and third stage matrices has $n \times k$ cross points. The total of these input and output stages combined is $2\,N/n$, yielding a number of cross points $(2N/n)(n \times k)$. The total number of cross points in the centre stages is $k(N/n)^2$ so that altogether the total number of cross points:

$$C = n(N/n)k + k(N/n)^2 + n(N/n)k = 2Nk + k(N/n)^2 \tag{4.1}$$

Now we shall determine how large C in Eq. 4.1 must be, if, like the single stage matrix, the three stage switch is to be non-blocking. We note, first of all, that to allow an outlet for each input of the first stage matrices, even if all the inputs are being used simultaneously, it is necessary but not sufficient that k be as large as n. To see that $k = n$ is not sufficient, let each of the input matrices be called X_i and let each of the output matrices be called Y_i as shown in *Figure 4.6*. Suppose that we want to make a connection from the input X_{11} of the first stage matrix X_1 to the output Y_{11} of the third stage matrix Y_1. We keep in mind that from each input matrix there is only one connection to each centre stage and there is only one connection from each centre stage to each output matrix. Now consider that all $n-1$ inputs, other than X_{11} of matrix X_1 are in use and also all $n-1$ outputs, other than Y_{11} of matrix Y_1 are also in use. Then the lines from the input matrix to the centre matrices use up $n-1$ centre matrices and also $n-1$ centre matrices are used up by the connections from the centre matrices to the output matrices. Finally, as a worst case condition, let us allow for the possibility that these two groups, each of $n-1$ centre matrices are mutually exclusive as shown in the figure. Then we have so far used $2(n-1)$ centre matrices. If we are still able to make a connection from X_{11} to Y_{11} there must be one more centre matrix available. Hence, in total we need a number k of centre matrices given by $k = 2(n-1) + 1 = 2n - 1$. Using the value of k from Eq. 4.1 we have:

$$C = 2N(2n-1) + (2n-1)(N/n)^2 \tag{4.2}$$

We find the value of n, for which C is a minimum, by the usual procedure of determining the value of n for which $dC/dn = 0$. If $N >> 1$, then, as can be verified, C is minimum when $n \approx \sqrt{(N/2)}$ and C becomes:

$$C(min) \approx 4N\sqrt{(N/2)} + 2N^2/\sqrt{(N/2)} \approx 4\sqrt{2}N^{\frac{3}{2}} \tag{4.3}$$

We have seen that the number of cross points for a single-stage switching matrix to connect N inlets to N outlets is $C(s.s.) = N^2$. Hence from Eq. 4.3:

$$\lambda \equiv \frac{C(min, 3stage)}{C(s.s.)} = \frac{4\sqrt{2}N^{\frac{3}{2}}}{N^2} \approx \frac{4\sqrt{2}}{\sqrt{N}} \tag{4.4}$$

for large N. For $N = 128$ the multistage switch requires half the cross-points, for $N = 2048$ the multistage switch requires an eighth of the points ($\lambda = 1/8$). The saving in cross points becoming more pronounced with increasing N. Still with such savings the number of cross points can become prohibitively large for large N. Thus, for $N = 256^2 = 65,536$, Eq. 4.3 yields $C(min) \approx 95 \times 10^6$. Hence, when a very large number of lines must be accommodated, switching structures are used which have more stages, even up to eight stages.

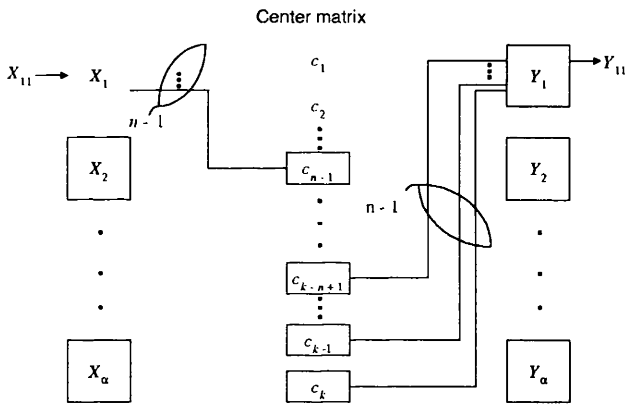

Figure 4.6: Three-stage switching matrix showing $k-n+1 \geq n$ to insure a connection between X_{11} and Y_{11}.

We have noted earlier that as a practical matter it is both reasonable and necessary to accept occasional blocking of a connection. Allowing such blocking will allow us to make further reductions in the cross point number. We have seen that in the three stage switching matrix we require that $k = 2n-1$ to guarantee no blocking. Let us inquire into the extend to which k can be reduced without introducing unacceptable blocking and to the extend to which cross points will be saved thereby.

Suppose we have made observations on the demands on the switching structure over an extended period and we have determined that an arbitrary selected

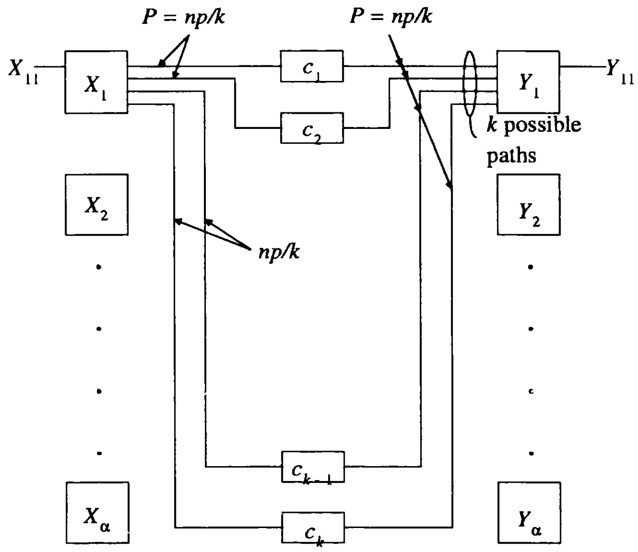

Figure 4.7: The k paths between X_i Y_i (illustration for $i = j = 1$).

input terminal is in use p percent of the time. Of course it is to be understood that the reason the terminal is in use is that there is a completed connection to some output terminal and there is presumably a data transfer in progress over some link joining an input and an output terminal. We turn our attention to the X_1 matrix shown in Figure 4.7. If only one of the n inputs is in use, then only one of the k outputs must be in use. Since there are k outputs the probability that a particular output is in use is p/k. Since there are n inputs, each with occupancy probability p, the probability that a particular one of the k paths from input to centre stage is in use is $P = np/k$, providing, of course, that $P \leq 1$. If $pn/k > 1, P = 1$. The probability that the path is available is $1 - P$. Similarly the probability that a particular path from centre stage to output stage is available is $1 - P$. The probability that both of these particular paths are available, thereby forming a particular idle link from input to output is $(1 - P)^2$. The probability that this particular link is busy is $[1 - (1 - P)^2]$. Now let us select one input X_{11} and one output Y_{11} in Figure 4.7. Server S is connected to X_{11} and wants to make a data transfer to set-top box T which is connected to Y_{11}. The end terminals having been selected there are only k links between them. The probability that all k links are busy is the probability B that the attempt of S to reach T will be blocked, and is then given by:

$$B = [1 - (1 - P)^2]^k \tag{4.5}$$

For a three-stage symmetrical switching structure with $N = 128$, with 16 first and third stage switches, the non-blocking structure would require 7680 points. This is derived from :

$$n = N/16 = 128/16 = 8$$

The non-blocking condition is $k = 2n - 1 = 15$. Using 4.1 we have:

$$C = 15(256 + 16^2) = 7680$$

cross points

To reduce the number of cross points to 7680/3 we must reduce k from 15 to 5. Using 4.5 we find, with $P = np/k = 8(0.1)/5 = 0.16$,

$$B = [1 - (1 - 0.16)^2]^5 = 0.002 = 0.2\%$$

Such a multistage switch could be used to replace the space switch we used in the simple design. The switch will connect users to video segments. The switch inputs for example can be replaced by the movie segments. The difference now will be the multiple access ability. Each movie segment should be accessible by a number of users simultaneously. Thus the last stage of the switch, the point where the user connections are connected to the required segments remains the same as before. Intermediate switching stages may be used to reduce switch complexity, but at the final stage we still need the control mechanism to transfer connections from segment to segment as instructed by the user. We can reduce switch complexity even further, since a number of users can be connected to the same segment, an identical path up to a point may be used, in the intermediate switching stages. Here we should point out that the switch should not abruptly terminate connections. Once a connection has been established it should not be terminated by a system management decision unless absolutely necessary. The multistage switch described can be rearrangeable under connection completion guarantees and the blocking probability arises on connection request where some part of the link may already be occupied by another user.

To reduce complexity at the final switching stage, where users are connected to segments, we propose an alternative design for this part of the switch. In this design a number of segments is grouped together, see *Figure 4.8.* Now the control over the access of the movie is divided. Some of the control is performed by the switch connecting to the groups of segments and some by the switches connecting to the particular segments. For each customer line, the customer channel control messages are split at the switch controlling the access to the grouped segments and handled separately. More than one customer channel per customer line implies the use of digital technology and time switching as we shall explain in the next paragraphs. A link is now formed to the switch controlling the access to a particular segment. Each of these switches can handle simultaneously a certain number of inputs. But the inputs to the switch can access any part of the movie segments even simultaneously. So as long as customers are on the same final stage switch they can access any part of the movie segments assigned to the switch and there is no contention.

By having two layers of switches controlling the customer channel requests we reduce control complexity. So instead of having one controller handling the whole request, we have two controllers each handling part of the request. The computing load on each controller is reduced, since the first controller looks at which part of the movie the connection is destined to, and the second controller connects to the exact frame required. The end controller will handle the connection for the time it is connected to the segments it is controlling. But there is a trade off here. There will be communications load from one controller to the other. The communications overhead depends on the rate the customer control messages are issued and the number of connections handled by the controllers. This rate is quite important both for the control load and the communications overhead. Charging schemes could direct user interaction to actions that do not overload the network.

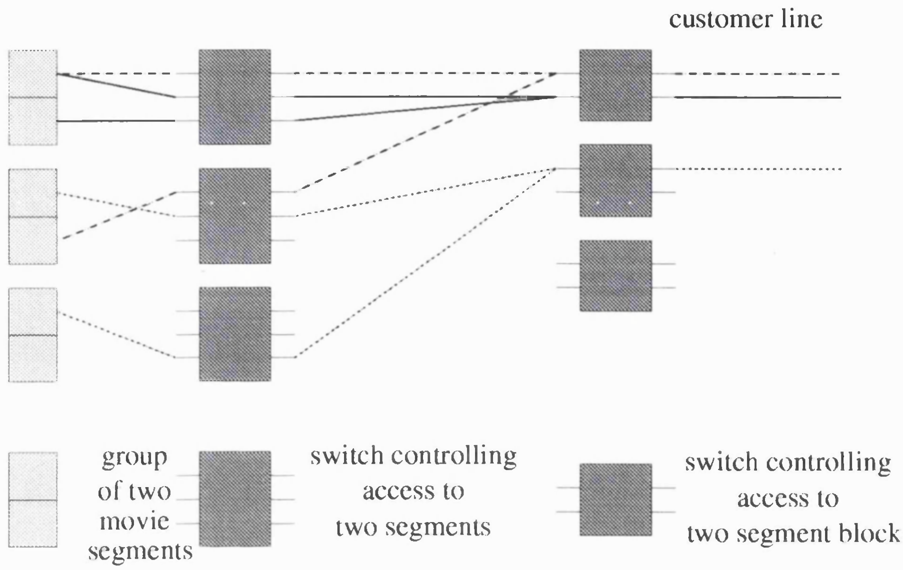

Figure 4.8: Representation of a split control switch

In *Figure 4.8* each customer line supports two customer channels. No more than three channels can access simultaneously the same group of segments. But once they have gained access to the switch controlling the group of segments they can access any segment.

The limitation in this design is that there is a blocking probability. Blocking arises when a new link needs to be established to the switch controlling access to a group of segments and there is no available switch input. Then the customer requests cannot be handled until one switch input is freed. This will prevent the server from executing the customer control message and the service from being interactive VoD. In the best case the customer will have to tolerate some delay in the worst case the customer requirement will not be handled at all. The tolerable delay and blocking probability will depend on the VoD server design and the QoS contract between user and provider.

Blocking probability for a VoD service must be quite small, for example a call should not be blocked more than once in 50 trials. The number of the switches controlling the segment access inputs, should be such that there is always a free input. An estimate is needed of the access probability per segment at a certain time. The statistics on customer habits during movie watching will provide a fairly good estimate as to what this probability will be. As already stated, pause seems to be the most frequently requested control function [49], hence at least this should be provided. Extra capacity should be reserved in case of increased interactivity demand. Loss of control or even connection will not be acceptable and tolerable by the customer and will contribute to deteriorating QoS. The final number of the second layer switch inputs will be a compromise between cost and tolerable blocking probability. As blocking or even disconnecting during movie watching constitute very bad QoS, this blocking probability should be almost zero. This may lead to switch underutilisation. A policy of temporarily restricting the provision of user functionality may be adopted and negotiated with the end user.

4.5 Digital Implementation

The switch will be implemented in digital technology. Some of the reasons are explained below. If we want to increase the number of customer channels per customer line we have to use digital technology. So for our application, we are forced to use digital technology as telecommunications operators did in order to realise huge space switches. A big spatial switch is often broken into a multistage switch which is essentially a network of interleaved time and space switches.

In the digital domain, the analogue signal is quantized and then transformed into sequences of binary system numbers, *i.e.* ones and zeros, also known as bits. Now the signal is transmitted as a sequence of bits. Any action on the signal, like transmission, multiplexing *etc.* is done in terms of those bits. The basic function of the time switch is the reordering of the signal frames.

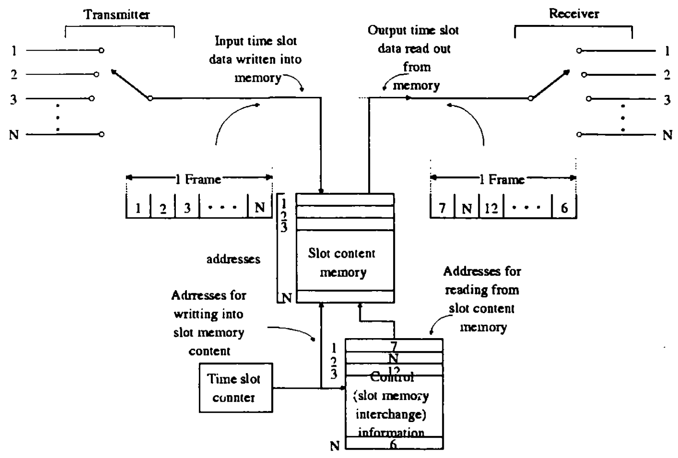

Figure 4.9: Time switch operation

A digital time switch works as follows. At the transmitting side, a multiplexer and an AD converter is needed, it multiplexes and encodes signals from N subscribers. At the receiving end the inverse process is carried out. The multiplexer and demultiplexer switches maintain a fixed synchronism so that the switches contact terminals 1 at the same time, terminals 2 at the same time etc. If there were a direct connection (as indicated by the dashed line), the single channel shown would carry N signals but there would be no provision for switching. To provide for switching, a slot-content memory is provided. This memory has N word locations, *i.e.* as there are time slots, and each location can accommodate all the bits in the time slot. The time slot counter runs at the time slot rate, that is, it increments its count by one at the end of each time slot. This counter provides the address at which a memory input word can be written into the memory. Accordingly we can arrange that the content of slot 1 will be written into memory location 1, the content of slot 2 in location 2, *etc.*, that each successive slot contents are written sequentially into the memory. (Actually to write a word into the memory it is generally necessary that all the bits of the word be available in parallel, that is, all at the same time). *Figure 4.9*, indicates that the bits in the time slot are available only serially, *i.e.* in sequence, one at a time. Accordingly it will be necessary to interpose, before the memory, a piece of hardware to convert from serial to parallel. A shift register driven by a clock at the bit rate can serve for this purpose. For the sake of simplicity we have not included such provision in *Figure 4.9.*

There is a second memory in the system. This memory is the control (slot interchange information) memory and it holds not the content of the slots, but holds addresses to the slot content memory form which words are to be read. The control memory also has N word locations and a word needs to accommodate as many bits as needed to write the number N. The time-slot counter which provides writing addresses for the slot-content memory provides reading addresses for the control memory. We can write address 7 in memory location 1 of the control memory. Accordingly, during the first time slot of the frame, the content of slot 1 will be written into the slot-content memory but during this first time the word in memory location 7 will be read from the memory. This is the word content of slot 7. Hence, altogether we have interchanged the contents of slots 1 and 7. Thus we have arranged that transmitting subscriber 7 has been switched to receiving subscriber 1. The same can happen for the rest of the connections.

This type of switch is known as TST (Time-Space-Time) switch and can perform in exactly the same way as the space switch considered previously, but offering higher accessibility. The space switch in the centre must switch channels between each frame. This means it has to operate at the slot transmission bit rate. This speed requirement is the effective electronic bottle neck of the current networks.

Time Division Multiplexing (TDM) drastically reduces the complexity of the switching units. Each switch only needs buffer memory to hold the address of the outgoing sequence. However the speed of operation of the space switching components is increased drastically.

This is how a network of time and space switches operates. In our conceptual model we have a space switch connecting the customer outputs to the correct segment of the movie. This would work as a real system as well but only for a

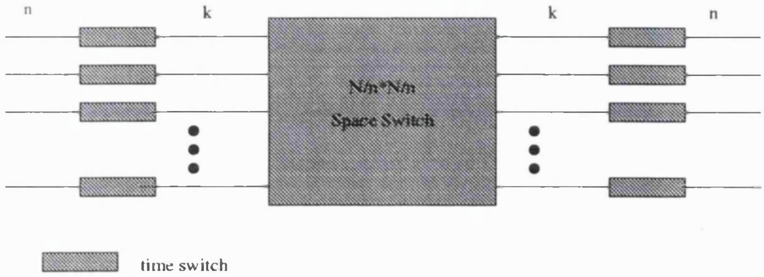

Figure 4.10: A TST switch schematic representation

limited number of customers; because as explained before the complexity of the space switches increases considerably with their size. In networks space switches have become multistage switches and eventually TST (*see Figure 4.10*) switches and other combinations (STS, TSSSST, *etc.*). For the VoD application we have divided the control between two layers of switches and we have incorporated time switches to handle more customer channels. This is the kind of switch most appropriate for the application. So for a real world application the original space switch will be replaced by a version of a TST switch.

Since the service allows multiple access, it can be considered a multicast process at that stage, at the final stage the blocking probability is reduced, blocking is encountered only on the first switching stage. By careful design of the VoD server we can reduce the blocking probability to nearly zero. This point is worth careful consideration. Serving a large customer base would be the real advantage of our design. Most designs already proposed in the literature [50], can serve a small number of users, something that makes them unsuitable for a broad customer base application. Their primary purpose was to test user reactions and possible usage patterns.

Our version of the TST switch would work as follows for the VoD application. The first stage time switches would be replaced by buffers, since we are going to load them only once. So that our original segments loaded in the magnetic disks become buffers. At the other end of the switch we have the customer lines. Each customer line supports a number of customer access channels (CACs).

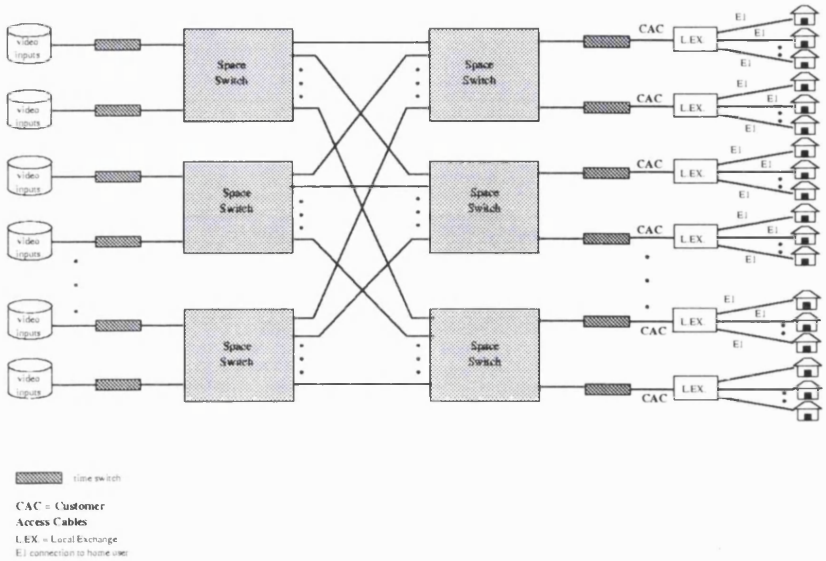

Figure 4.11: System diagram

Between the buffers and the customer lines we have the two layers of the space switches. At each customer line we have a digital time switch. Space switching is directing customer channels to the ordered segment and time switching is ordering the retrieved information to the correct customer channels. The feature we need to add is the control mechanism between the two switch layers. This is the heart of our design shown in *Figure 4.11.*

The space switches connect every customer output line to the correct segment at the correct time. This is done by a programme. Actually as we mentioned earlier it is two programmes working in conjunction. They should be capable of processing the CACs in parallel. Parallelism here is a very important characteristic. If we cannot cater for all the customer channels simultaneously, then our service is no longer interactive VoD. Each customer output line would need its own control in parallel with the other lines.

But the order of the frames after the first layer space switches will not necessarily be the correct one. Again time switches in the customer output lines would reorder the frames. In this way the transmitted signal will have the correct frames in the correct order for the correct customers.

The control structure is the most important part of our design. The software controlling switch operation does not exist. As a computing task it would require substantial processing power, for a large sized switch. It should be able to handle the CACs in parallel and we need at least several thousands of customer channels running in parallel. To carry out the task we would need a very fast, possibly parallel computer, with a single or multiple processor. Each user channel would effectively require its own processor. The control computation is not a simple one, but the advancements in computer processing power of the past decade will make this task easy to realise.

Implementing parallel control has its advantages and disadvantages. With parallel control all the customer control messages are processed independently. With this method, we avoid the bottlenecks created due to sequential computing. Any number of control messages will not create congestion on the control mechanism. The main disadvantage is the more complex software design requirements and the amount of processing power necessary to complete the task.

We could assign one processor per customer input line. Then this processor will have to execute the commands of all the user channels, say x. If it takes y milliseconds to execute a command from the moment it is received then if all the users issue a command it will take $x \times y$ milliseconds to be executed, this gives a waiting time of xy seconds.

Today's high-end parallel computers can support speeds of GFLOPS (10^9 floating point operations per second) such as the Cray 2000 [51] or even TFLOPS (10^{12} floating point operations per second) such as the Intel ASCI Option Red Supercomputer [52]. Parallel computers are still expensive since their use is limited, as well as, applications supported by them, but prices will drop once popular applications exist.

Alternatively we could use one single processor to handle all the customer lines in parallel, and for each line process in parallel the customer channels. The choice of processor depends on cost and availability, but it should be borne in mind that the cost of faster processors increases far more steeply than the performance they offer.

Our design approach lies on the computational side of the problem. This has the advantage that we are using already existing hardware which we adapt to perform something similar to its original function by adding a different control method. This approach I think is the most cost effective one, since we do not need to construct new hardware to implement it, nor any changes or adjustments are needed to the existing telecommunications network. Also the number of served users is quite large and random access is allowed. This makes our design an ideal platform for further multimedia applications. How difficult it is to construct the control method proposed here is the topic of later chapters.

Future development of the design would be on the control and design side of the problem. On the control side is the development of the suitable software. On the design side the number of switches and inputs to switches has to be specified. Finally, the cost per customer will be the determining factor for the implementation of the application.

4.5.1 Pricing memory and controller-switches

Taking into account the huge compression ratio achieved by MPEG-1 of 200:1 the segment design can be as follows. The MPEG-1 version of the movie which in terms of memory will be around 1 GB for 90 minutes should be loaded onto fast access memory, such as RAM and DRAM (or hard disks with SCSI controllers). One second of the movie 0.18 Mbytes of memory. With today's prices (December 1999) 64 Mbytes of RAM cost 40 pounds; 5.9 minutes of video could be stored in 64 MB of RAM. So the price of a memory system based on 64 MB RAM SIMMs would cost 600 pounds for one movie. For a server with 5 movies capacity this would amount to 3,000 pounds. We also have to take into account the drop in prices when these elements will be mass produced. For our purposes though we shall use segments of less capacity, this will be of one second of video for reasons explained in the next pages.

The cost of the space switch and parallel controllers needed should be added to the estimated price.

4.5.2 The controller functions

We have a two level space switch. One level controls the access to the exact segment and the other the access to the group of segments where the required segment belongs. By having two levels of switching we reduce the complexity of the switch, but do introduce the extra cost of communication between the two levels. The blocks needed on each level are to be calculated based on the desired blocking probability, the predicted user access taking into account the interactive capability the user should have over the delivered material (allowing for interactive functions) and the cost of their production. The server should be able to cope with the worst case scenario, which is prime time demand and provide QoS to the customer within the allowed limits as set by the provider (both content and network providers) and the customer. In a 100,000 subscriber area (here we do not take into account the market penetration or service take-up rate), at prime time i.e. between 8-11pm, (AT&T report suggests) 40% of the customers will want to access the system. This means that 40,000 customers, out of a 100,000 customer base, will require to receive their requested choice of film (or any other video material). The blocking probability should be as low as possible, a realistic value is this of 0.1% in the final attempt, one and a half nights per year the user will not be able to access the service. This means that only 40 users could be denied service and the system should consequently have been built in such a way as to withstand this demand.

4.5.3 VCR function provision

A request is issued by the user for a VCR function such as fast-forward, rewind and random access in either direction. At first the request from the first stage controller is forwarded to the second stage controller. From then, we have two courses of action. If the action takes place within the segments of the group no more communication is needed. In the second case the action exceeds the segments within the block. In this case the first stage controller will have to move the connection to another second stage controller. For this to be done the second stage controller should send a signal to the first stage controller when the connection is at the last or first segment, so that the first stage controller can request and if successful reserve a place on the next stage to ensure uninterrupted viewing. There is a need to use pre-empting to ensure smooth viewing.

Here arises again the blocking problem. As we stated the actions of one user should not limit the film delivery to the other users. If the required second stage controller works on full capacity, then there is no place for the new stream and consequently the user will be blocked. We have to ensure that at peak time there is always some spare capacity available on the second stage controllers in order to accommodate the extra requirements of interactive functionality which should be available to the users. A way of easing the problem would be to move to the next block, if empty, instead of blocking the connection. How big a jump is from one block to the next, depends on the number of segments per block. On this number of segments per block depends on the number of second stage controllers required and their capacity.

For example if the 40,000 requests are uniformly dispersed in 15 minutes, we have 44 requests per second. If the group of segments comprises 5 one-second segments, we have 220 users per group of segments. If a generous 5% uniformly distributed, is likely to request a VCR function, this means that the second stage controller should be able to handle another 11 users giving a total of 231 users.

The server designer needs to take into account these considerations under the scope of cost and Quality-of-Service provision. Availability should be provided, but not under prohibitive cost.

Controllers and access to the stream composition.

Now if each segment is viewed by 44 users simultaneously, the server should copy or multicast the same segments to all 44 output streams at the normal viewing rate. To provide fast-forward-reverse viewing, the parallel control, should be able to run a stream independently through the segment but at a faster rate. For example, as already mentioned, only the I intraframes positioned at the beginning of the segments, could be delivered to the end user.

4.6 Summary

In this chapter we have presented our design for a VoD server, in a distributed environment. The server is based on using independent memory units, which we call segments. A two-tier control layer will provide the material to the end user. The segmentation process offers a number of advantages, such as multiple access by a number of users, limited only by the memory characteristics, easy random access to any part of the material. The smaller the segment size the less the granularity of these actions. Segment size selection can be done according to material popularity and cost.

A real world switch would be a digital multistage switch, a mixture of time and space switches, as most modern switches are. The difference will be at the input stage, which is not customers, but rather movie segments. The control mechanism which will transfer connections from segment to segment according to user orders, can be done in two stages to reduce complexity.

All major computer manufactures are interested in VoD server design and a number of servers has been put forward. Most of these servers though can only support a small customer base and are very expensive.

In the next chapter we shall present a review on the existing bibliography on VoD systems.

Chapter 5

Review of bibliography on VoD systems

5.1 Introduction

As already stated in the previous chapter we propose a distributed network of VoD servers as part of service implementation. The networked system can have centralised or distributed control. In the centralised case a unique manager communicates with all the servers and decides where and when to realise connections, which movies to load and other system management functions. The manager has knowledge of the total system state. In the distributed control case no unique manager exists, but rather control is distributed to the individual servers. Each server should manage its own resources based on a number of criteria.

Each process has its advantages and disadvantages, depending on the particular application. First we shall present the general case for centralised and distributed management and then we shall discuss how it applies to the specific application of VoD and review a number of papers which examine different aspects of the distributed solution for VoD.

5.2 Centralised and distributed management

The argument of centralisation versus distribution, is not a new one. It has been around ever since distributed systems came to existence. The arguments are more or less known to all involved and are primarily concerned with size of structure, expandability, scalability and controller power. The basic arguments are:

- **expandability**: adding a few more units to the existing network is much more cost effective than replacing the old centralised machine with a new one. We can also use the existing equipment and expand the structure indefinitely or as much as the system design allows.
- **cost/performance**: as the power to cost ratio increases considerably for more powerful machines, a network of many not so powerful machines becomes an attractive, less expensive alternative.
- **availability**: since distributed systems usually replicate data and have built-in redundancy in all resources that can fail, distributed systems have the potential to be available even when an arbitrary number of failures (single-point) occur.
- **scalability**: the capacity of any centralised component in a system imposes a limit in the maximum size of the structure. Ideally distributed systems have no centralised components, so that this restriction on the maximum size of the system does not exist. Naturally there can be many other factors that restrict system scalability, but distributed system designers are doing their best to choose algorithms that scale to very large numbers of components and consequently lead to future proof systems.
- **reliability**: availability is but one aspect of reliability. A reliable system must not only be available, but it must meet its performance requirements, even when component failures occur. The algorithms used in distributed

systems must not only behave well when the underlying virtual machine functions correctly, but they must also be capable of recovering from failures in the underlying machine environment and work with the necessary components, thus isolating the malfunctioning ones [53] page 5.

- **performance considerations (parallelism)**: in a distributed system we have an increase in performance through a high degree of parallelism [54] page 12. Actions can be performed in parallel where possible, adding speed and efficiency to system performance.

- **management and control** requirements: in a centralised system management is simple since all actions and decisions are taken at one point and then performed at the periphery. Managing a distributed system is more complex. Distribution of actions requires distribution of management. Every unit should be able to take its own management decisions based preferably, on the rest of the system. This aspect of distributed systems constitutes, their main source of complexity.

The above advantages make the distribution of data in a network a very attractive approach to the efficient and cost effective provision of information and system organisation. Also, it can potentially expand to cover future needs, growing as the system grows. The provision of information through new multimedia services, is an area which, as we propose, could greatly benefit from component distribution.

5.2.1 Problems of Distribution

The main reason that the design of distributed systems is so hard to achieve is the enormous complexity of these systems, which is not yet fully understood. However, we now understand many aspects of distributed systems and many sources of complexity.

To understand the source of complexity in the distributed systems, we can start by comparing distributed systems to the railway system, another "distributed system" most people are familiar with. The railway system as we know it today has taken a century and a half to develop into a reliable and safe transport service. It took time before the complexity of the railway system was understood. The development of its safety has been at the cost of errors and mistakes that have cost money and lives. In this one and a half century the railway system has grown and expanded in spatial terms and managed to cover almost all of the developed world. Along with it, the design and safety requirements have grown as well. More lines and new destinations meant new problems for the network designers and engineers. These problems had to be solved along the way.

The design of distributed computer systems has only been developed for around two decades or so. Designing a sizeable distributed computer system, will probably be at least as difficult as designing the railway system. One factor that the complexity of the system depends upon is its intelligence. Although a strict, universally acceptable definition of intelligence does not exist, it could be argued that it is associated with learning from and adapting to the environment. Possibly adaptation and conditioning have nothing to do with intelligence but they can produce a seemingly intelligent behaviour which is all we are trying to achieve [55] page 31.

Another source of complexity in distributed systems is that the connection of well-understood components may generate new problems not apparent in the components themselves [53] page 9. Additionally, we have to take into consideration how more than one system units co-operate and what system organisation we need, to avoid chaos and inoperability.

Yet another source of complexity is the organisation and administration of the system. The definition of a distributed system may vary from author to author and there is no absolute definition. The definition employed here, is that a system made up of untis which base their actions on information received from the rest of the system. The distribution of control to the components is very important. For a system to be distributed the control needs to be distributed. To achieve this, a control algorithm that performs acceptably and keeps the system within operational bands is needed. This is the main reason why a distributed system is much more complex and harder to design than a centralised one. To design and optimise this control algorithm and make it a reliable one as well, may prove to be one of the most expensive and time-consuming requirements in building a distributed system.

This is also the topic we have chosen to examine more closely. In a database environment this can be translated into building an algorithm that controls and administrates the information distribution in the database network. This is applicable for any information providing service as well.

5.2.2 Optimisation and Automation of distributed control

A few years ago information distribution and organisation could be done manually. The system administrator was responsible for distributing information across the network. In order to do this he had to assess the frequency a certain piece of information was accessed at a certain location and accordingly decide if it was worth while placing it there. This procedure should be done at regular intervals since information demand is a dynamic process. This is not the best way of information distribution. It is prone to human error, not very efficient, time consuming, costly and does not scale up with huge increase in information demand.

A way to overcome the problems of human data organisation would be to create an automatic method, just like machines took over many manual jobs. In this case it will not be a machine but an algorithm that will do the task. Such an algorithm should automatically distribute data in the system without any external intervention. The distribution algorithms should be adaptive–take feedback from the network–and be able to decide based on this feedback and other considerations on how to distribute the data at any instance. Additionally such an algorithm should exploit in the best possible way the available system resources.

System resources usage is the communication overhead due to redistribution and access to data, the traffic overhead due to network usage, the processing power required for each decision, the amount of memory occupied by the data, the cost of components for system repairs and extensions, company organisation, *etc.* All system resources may not have the same importance in the operation of the system, for example in a system where processing power is very limited, and therefore very expensive, everything should be done to use it as little as possible, or if the system memory is near to saturation any action that involves using it should be avoided, unless more memory is added. The distribution algorithm should be able to evaluate the importance of each of the system resources and adjust its performance, according to the current situation in the system. It should also combine all the different system resources and their importance. This will give it the power to decide how to distribute the data in the most efficient and cost effective way and thus use the system resources in the most efficient manner.

5.2.3 Distributed method optimisation

What we described above fits the description of an optimisation process. Our algorithm could either be able to optimise itself or it should be optimised when we apply it to the system. In the later case the algorithm should be optimised in a near to real world system and tested before application. A number of optimisation methods exist, a few will be reviewed later on.

Before we go on with the description of the optimisation methods, let us first state more precisely what we mean by optimisation. The conventional view is presented well by Beightler, and Widle (1971, p1):

> *Man's longing for perfection finds expression in the theory of optimisation. It studies how to describe and attain what is Best, once one knows how to measure and alter what is Good or Bad....*

Optimisation theory encompasses the quantitative study of optima and methods for finding them [56] page 6.

Thus optimisation seeks to improve performance towards some optimal point or points. Note that this definition has two parts: (1) we seek improvement to approach some (2) optimal point. There is a clear distinction between the process of improvement and the destination or optimum itself. Yet often optimisation procedures focus solely upon convergence.

Although this is the strict definition of optimisation, in this project we are not seeking the absolute optimal point as already mentioned, a better performance beyond what is already available would often suffice.

There exist four main types of optimisation methods. They differ on the way a method is employed to approach the optimal point. The four most popular optimisation methods are the calculus based, enumerative, random and randomised.

The calculus based methods have been studied heavily. These subdivide into two main classes: indirect and direct. Indirect methods seek local extrema by solving the usually non linear set of equations resulting from setting the gradient of the objective function equal to zero. This is the multidimensional generalisation of the elementary calculus notion of extreme points as illustrated in *Figure 5.1.* Given a smooth unconstrained function, finding a possible peak starts by restricting the search to those points with slopes of zero in all directions. On the other hand, direct (search) methods seek local optima by hopping on the function and moving in a direction related to the local gradient. This is simply the notion of hill-climbing: to find the local best, climb the function in the steepest permissible direction. While both of these calculus based methods have been improved, extended, hashed and rehashed, some simple reasoning shows their luck of robustness.

First, both methods are local in scope; the optima they seek are the best in the neighbourhood of the current point. For example suppose that we start the

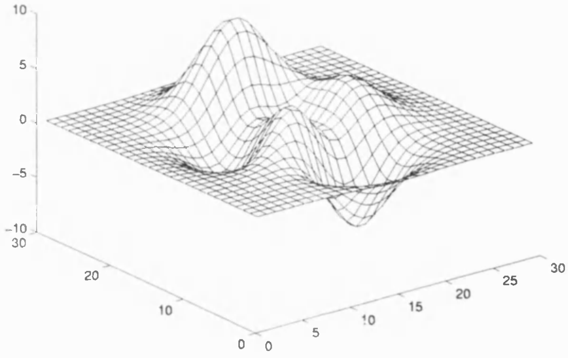

Figure 5.1: A multiple-peak function causes a dilemma. Which hill should we climb?

optimisation process near one of the lower peaks in *figure 5.1*. Clearly, starting the search or zero finding procedures in the neighbourhood of the lower peak will cause us to lose the main event (the higher peak). Furthermore once the lowest peak has been reached we cannot be certain that this is the global maximum. If the second derivative cannot be defined, further improvement must be sought through random restart or other trickery. Secondly, calculus based methods depend upon the existence of derivatives (well defined slope values) and analytic functions. Even if we allow numerical approximation of derivatives, this is a severe shortcoming. Many practical parameter spaces have little respect for the notion of the derivative and the smoothness it implies. Theorists interested in optimisation have been too willing to accept the ideal mathematical world painted by the great eighteenth and nineteenth century mathematicians. Quadratic objective functions, ideal constraints, and ever present derivatives are only a very small part of the real mathematical world. The real world of search is fraught with discontinuities and vast multimodal, noisy search spaces as depicted in a less calculus friendly function in *figure 5.2*. It comes as no surprise that methods depending upon the restrictive requirements of continuity and derivative existence are unsuitable for but a very limited problem domain. For this reason and because of their inherently local scope, we must reject calculus-based methods. They are insufficiently robust in unintended domains.

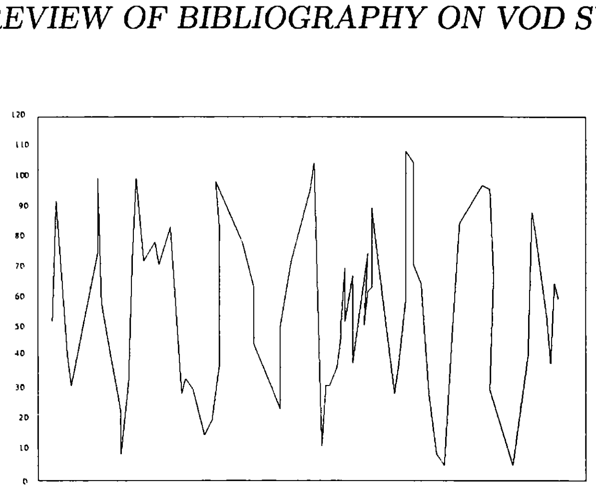

Figure 5.2: An example of a less calculus friendly function

Enumerative schemes have been considered in many shapes and sizes. The idea is fairly straightforward; within a finite search space, or a discretised infinite search space, the search algorithm starts looking at objective function values at every point in the space, one at a time. Although the simplicity of this type of algorithm is attractive, and enumeration is a very human kind of search (when the number of possibilities is small), such schemes must ultimately be discounted in the robustness race for one simple reason: lack of efficiency. Many practical spaces are simply too large to search one at a time and still have a chance of using the information to some practical end. Even the highly touted enumerative scheme *dynamic programming* breaks down on problems of moderate size and complexity, suffering from a malady melodramatically labelled the "curse of dimensionality" by its creator (Bellman, 1961). We must conclude that less clever enumerative schemes are similarly, and more abundantly, cursed for real problems.

Random search algorithms have achieved increasing popularity as researchers have recognised the shortcomings of calculus based and enumerative schemes. Yet random walks and random schemes that search and save the best must also be discounted because of the efficiency requirement. Random searches in the long run, can be expected to do no better than enumerative schemes. In our haste to discount strictly random search methods, we must be careful to separate them from randomised techniques. The *genetic algorithm* is an example of a search procedure that uses random choice as a tool to guide a highly exploitive search through a coding of parameter space. Using random choice as a tool in a directed search process seems strange at first, but nature contains many examples. Another currently popular search technique, *simulated annealing,* uses random processes to help guide its form of search for minimal energy states.

The important thing to recognise at this juncture is that a randomised search does not imply a directionless search [56] page 5.

Optimisation is highly desirable in all systems, but can prove quite costly. In our distributed VoD server system, we shall try optimise use of system resources, applying a management system. We shall try to make this system as simple as possible while still providing efficient system management.

5.3 Review of bibliography

Due to the short history of the field, only a small number of service design proposals have been put forward. A few of the most important ones will be reviewed. All four works presented next, argue for the benefits of a distributed solution, against a centralised one and give support to our proposal for a distributed type of solution.

5.3.1 A store-and-forward architecture for video-on-demand service [1]

The most important components of this architecture are: the information providers or information warehouses (IPw), central office (CO) service circuits, and the customer premises equipment. The information providers are geographically distributed entities responsible for long-and medium-term information storage. They are also capable of dispensing information in parts, or in total, via high speed dedicated lines that connect them to the network central offices.

Video and other information programmes owned by vendors are stored in IPws using various storage media: optical, magnetic and electronic, with a hierarchy of access mechanisms: non-random access for archival material, video juke-boxes for medium-speed access, magnetic disk drives for on-line popular material, etc. The information is stored digitally, in a compressed form.

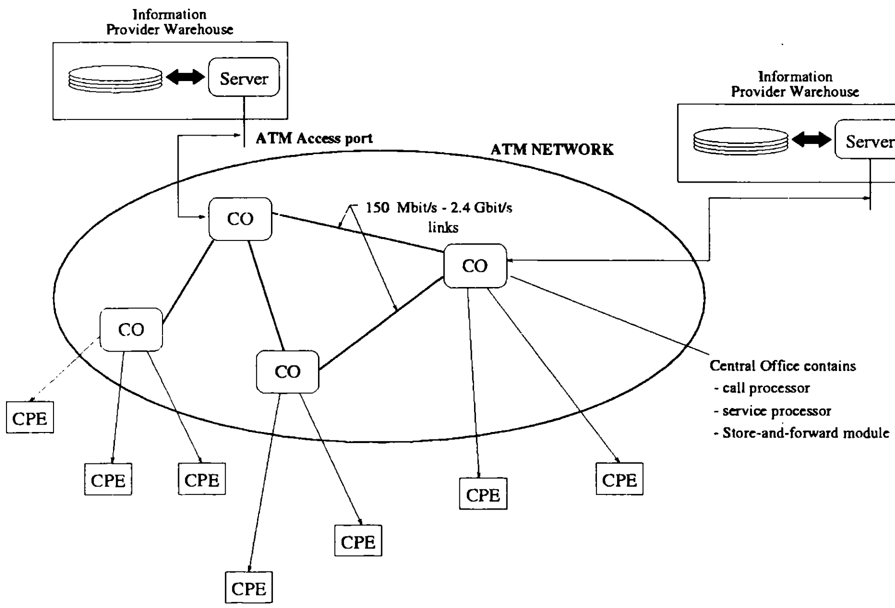

Figure 5.3: Store-and-forward Architecture.

Other than for the addition of the IPw, see *Figure 5.3*, this architecture represents a typical local access and transport area (LATA) network in a B-ISDN environment. For the VoD service, subscribers place requests with their local CO (local exchange) via user interfaces, using an upstream control channel. In turn, the CO places a request with an appropriate IPw.

The IPws are connected to the network via high-capacity STM-1 (155Mbit/s) or STM-4 (622Mbit/s) lines. Therefore, the information trucking between the IPw and the customer's CO can be done faster than real time (i.e. 1.5Mbit/s), and the requested programme can be delivered in segments. In order to take full advantage of this "speed-up" factor, the transfer rate of the stored material to the network should be equal to the network transmission rate; i.e. 150Mbit/s. The retrieval of the video segments is managed by an ambassador service processor IPw-SP. The ambassador processor, while possibly being located on the vendor's premises, co-operates with the network and possesses relevant knowledge of the network state. This knowledge allows the processor to effectively schedule the information transport to requesting COs, taking into consideration the time budget of each request and the traffic conditions in different parts of the network.

At the user's local CO, the information is buffered, the data-rate is then converted to the video coding rate (i.e. DS-1), and then possibly is decoded to the original video signal form (e.g. analogue). The video signal is then transported to the user in the form which corresponds to the local access video switching and transmission parameters and the user's CPE capabilities. In this paper, it is assumed that all users possess terminal equipment which accepts their compressed or non-compressed signals. The DS-1 video decoder may therefore be located at the CPE or be part of the service circuit pool at the CO. In either case, the matching of video rates would be done at the CO and the channel from the CO to the subscriber is used for real-time data transfer.

This architecture takes advantage of the high transfer rates of currently available and future, storage systems and of high-speed trunks, to provide the rate speed-up and consequently vendor and network resource sharing through time-multiplexing. The store-and-forward mechanism, by being more general than a real-time transfer system, allows alternative (to real-time) information delivery schemes. Thus, better utilisation of the network resources and improvement in the overall service economics are among the advantages of this architecture. The flexibility of the proposed technique allows one to employ, if necessary, higher layer protocols for error and/or cell loss control. By distributing the complexity of the supporting hardware and software infrastructure, the reliability of the service can be improved.

In this paper the virtues of a distributed VoD system are shown. The system places the material close to the user. The information providers may form a separate network connected to the service provider. The system has no loading control mechanism, although the distinction of popular to nonpopular material is made. It makes a good introduction to using a distributed delivery system for VoD, where information provider and service provider can be two different entities.

5.3.2 Content Hosting in Full Service Broadband Delivery infrastructures: Efficiencies, Architecture and Policy [2]

In this paper 4 different Architectures for VoD service are presented and compared. The comparison is done in relative terms of cost and efficiency. The designs start from a centralised approach to reach a centralised-distributed approach where the cache servers are used in order to satisfy the bulk of requests. The VoD servers are combined into an ATM/SONET network.

Architecture 1 illustrated in *Figure 5.4* is called a regional server approach, where the total population is divided into n regions, each region served by centralised video servers that operate exclusively within a region. Thus, all VoD requests are handled by the server pool that is connected to the hub switch of that region.

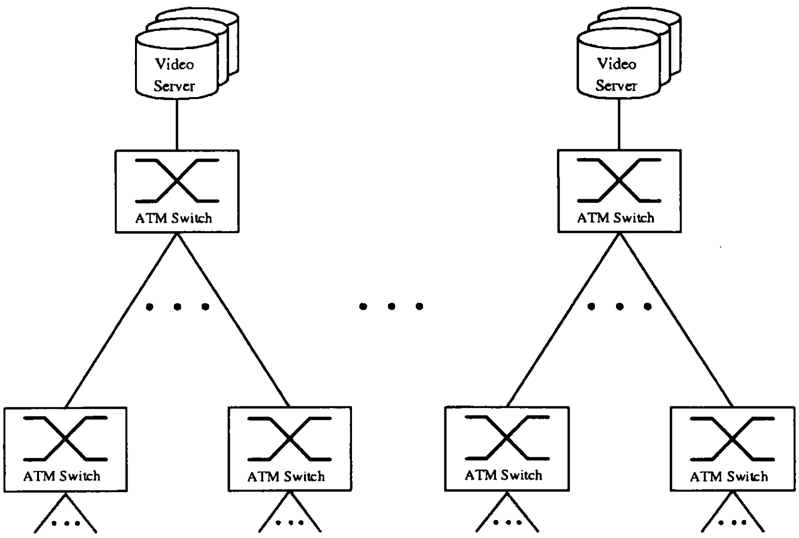

Figure 5.4: Architecture 1

In Architecture 2 shown in *Figure 5.5*, a variation on 1, subscriber traffic is divided between the edge switches and the (m) centralised hub switches across the (m) regions. For a given scenario this architecture has the same number of trunks as required for Architecture 1, and thus the relative economic efficiencies are the same.

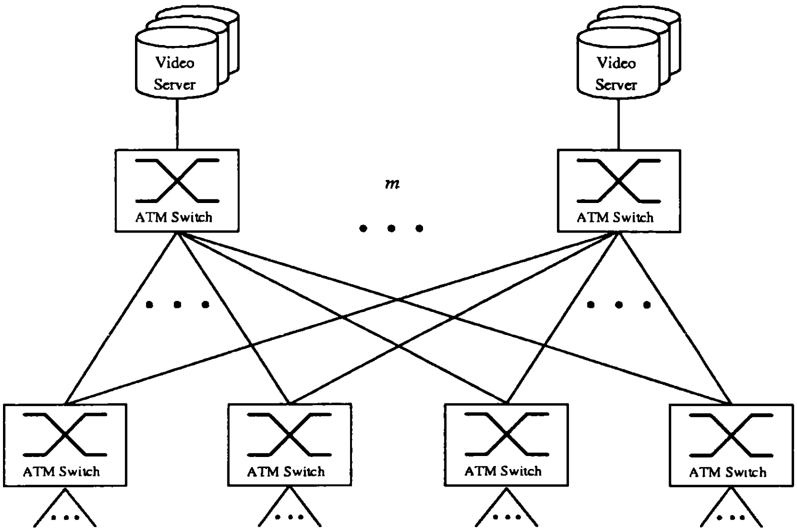

Figure 5.5: Architecture 2

The advantages of this architecture are the potential balancing of traffic and increased service availability.

In Architecture 3 shown in *Figure 5.6*, cache servers are placed at each of the edge switches. Depending on the titles placed in these cache servers, the number of VoD requests-that must be handled by the hub servers-can be greatly reduced, and consequently the number of SONET trunks between the edge switches and hub switch, as well as the size of the hub servers will be reduced. To make this scenario as effective as possible, the highest demand titles should be placed on the cache servers.

Architecture 4 is illustrated in *Figure 5.7*, it is essentially an extension of architecture 2 with the addition of cache servers. The same advantages of architecture 2 apply with the addition that the cache servers will minimise the demand on the inter-switch network (as in architecture 3) and potentially provide even greater improvements in service availability.

An important variable of this architecture is the title break point (TBP).

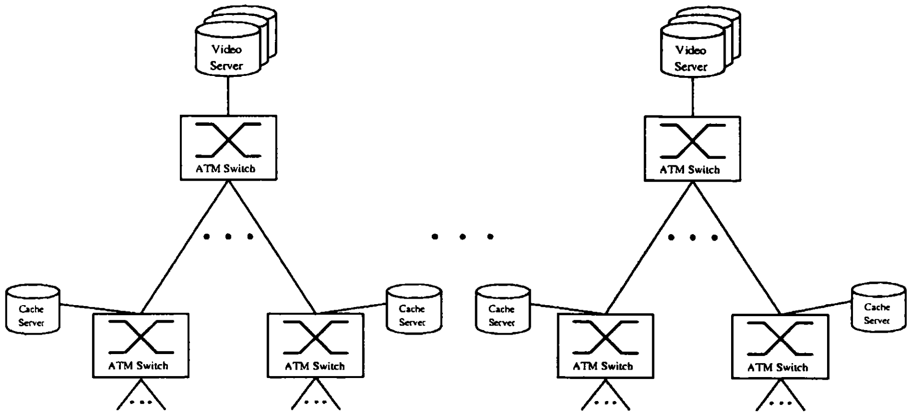

Figure 5.6: Architecture 3

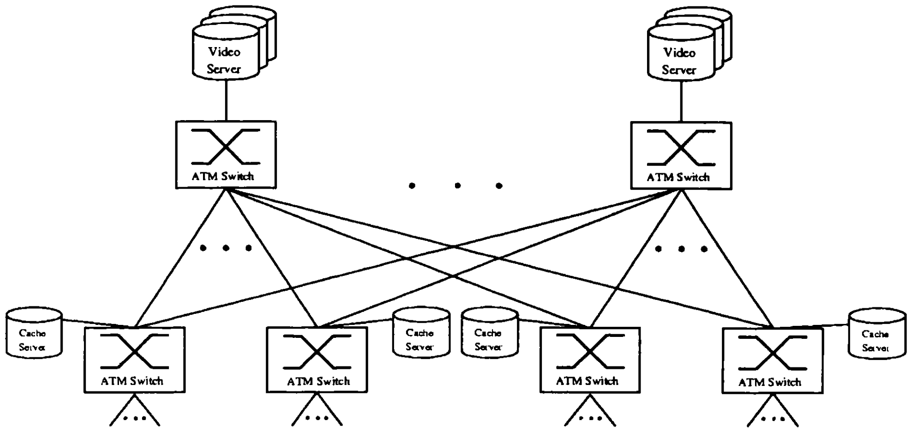

Figure 5.7: Architecture 4

Given a movie library of certain capacity and demand decreasing with increasing order of placement, the TBP divides the library in two, the more popular and less popular movies.

For their analysis they used a stochastic probability model, for the subscription rate, the call probability model, the stochastic title probability model and the video server capacity model.

Four different scenarios were analysed and compared. In all four, they used 1000 homes passed (HP). In the first an MPEG-2 bit rate of 4Mbit/s a subscription probability (ps) of 10% and a title library size of 1000. In the second the subscription probability was raised to 40%. The third and forth were the same as 1 and 2 but the MPEG-2 rate used was increased to 6Mbit/s. The probability of a subscriber requesting VoD service was held constant at 10%.

They also took into account a predicted fall in bandwidth and server prices. The ATM and SONET costs reduced by 5% per year, the server storage capacities increased by 27% per year, the disk throughput increased by 22% per year, and the storage costs reduced by 30% per year.

In all scenarios they calculated the potential savings against the centralised architecture with different title break points and initial population of 250k and 500k in the time periods of one and two years.

The title request probability was determined using a Zipf function, see sec. 8.2.1.

Summary of results.

Using the above mentioned scenarios in the four topologies examined, they came to the following findings:

1. Placing video servers near the edge switches provides economic efficiencies (up to 50% savings in total costs in certain cases) when compared to architecture 1 for a wide range of title break points, from 5% of the library attracting 56% of total demand to 50% of the library attracting 90% of total demand.

2. Economies of scale: cost per 1000HP drop when going from 250k HP to 500k HP

3. When demand is high (ps = 40%), library size has little effect. This is based on the assumption used in the model that the library size has already passed an initial threshold where an increase in the number of titles will not necessarily (significantly) increase demand. That is the fact that there are 500 or 1000 titles in a library will not affect a customer's probability of requesting VoD service.

4. Caching titles near the edge switches is more economical when predicting the decline in server costs for the server technology of tomorrow. SONET and ATM technology are more mature than present video server technology and their costs were expected to decline slightly over the study period. By taking industry trend curves, and driving future video server costs and capacities, the expectation is that server costs will decline relatively faster than ATM and SONET costs, thus make the caching scenario more attractive.

5. Caching looks more attractive for higher call probability. This is based on the fact that higher demand increased the economic savings of caching when ps increased from 10% to 40%, therefore an increase of call probability, pc = 10%, will have a similar effect.

6. MPEG bit rate seemed to have little effect on the relative advantages of caching, though in the light demand scenario with smaller total homes passed, caching could in fact have higher costs, depending upon the break point. Of course, total costs increased as the bit rate increased.

7. In the case of smaller HP, smaller subscription probability, higher bit rate and larger library size, the most efficient scenario was to use central video servers with no caching. The low demand of the low HP and low subscription probability do not create enough traffic to offset the increased costs of video servers in the caching scenario.

8. For increased demand and a higher number of homes passed, the differences in savings between break points was not significant, though the range of break point from 1 to 3 seemed to provide the highest expected average savings in most cases.

In this paper a regional centralised architecture is compared to a localised distributed one, as well as with variations such as interconnected regional servers and localised servers. The localised distributed architecture is the most cost effective one, when there is substantial demand for the service. They make the distinction between highly popular movies and less popular ones. All these are parameters, which are used in this work and will be presented. In the long term they predict that co-operation between service provider and information provider is essential for the good operation of the service.

5.3.3 A cost comparison of Distributed and Centralised Approaches to Video-on-Demand [3]

The storage costs for a centralised vs. a distributed VoD storage system are compared. They propose a VoD service design where a centralised VoD server is connected to a number of Front End Servers (FES). Depending on the FES capabilities of playing movies the system varies from centralised when no such capability is present, to distributed when the capability exists.

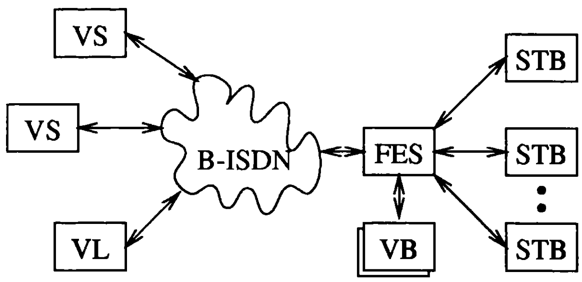

Figure 5.8: VoD with FES servers, VS denotes the video server, VB denotes video buffer, VL denotes video library

The difference according to the authors, between a centralised and distributed system relates to where the video streams are originally generated. Both approaches have a topology similar to the one depicted in *Figure 5.8*, but in the distributed case, many of the user requests will be satisfied by the video buffers located at the FES, without need for a connection to the VoD server. In the centralised approach all requests are passed from the front-end to a certain video server, increasing the bandwidth requirements in the core network. *Figure 5.9* represents the difference between a centralised and distributed system for an example viewer population. In this comparison they assume a total user

Table 5.1: Configuration for a Distributed System

% Requests Served by FES	80	85	90	95	99
Number of Movies Required	45	55	82	200	680
Ratio of VS:FES	1:5	1:6.7	1:10	1:20	1:100

population of 100,000 customers in one geographical area, with a peak activity of 40%. This gives a total active stream requirement of 40,000. It is assumed that 1000 movies must be available to all users. The analysis is easily repeated for other values of customer population, but they have found that this is a representative example based on the likely evolving network infrastructure and it gives representative results. Due to the algorithmic nature of the disk allocation optimisation, a general analytic result is difficult to obtain. Using an access probability function they calculate that on average only 82 titles are required to satisfy 90% of viewers requests. The probability function is based on Zip function (see sec.8.2.1) and empirical data too. Hence, in the situation in *Figure 5.9*, they are able to have ten front-end servers all connected to a single video server, which is responsible for serving the 10% of streams not satisfied at the FES. The choice to satisfy 90% of the requests at the FES is an arbitrary one and can be easily varied by changing the number of movies stored at each FES. This will also affect the number of FES's that can be connected to a single VS in order that the VS is of the same stream processing capability as the FES's. This is in the distributed system shown above, the FES's and VS are both capable of supporting 4000 concurrent streams. Whereas the equivalent centralised VS must support 40,000 streams and store 1000 movies. Table 5.1 shows how the parameters of the distributed system vary depending on the percentage of requests they aim to satisfy at the FES. Using the disk allocation strategy they developed they analysed the cost in both centralised and distributed cases.

From their results they conclude that the distributed approach operates more efficiently from a storage point of view, when the FES satisfies approximately 85% of requests. That is, the FES should store about 55 movies given the parameters used in their example. They also conclude that the distributed

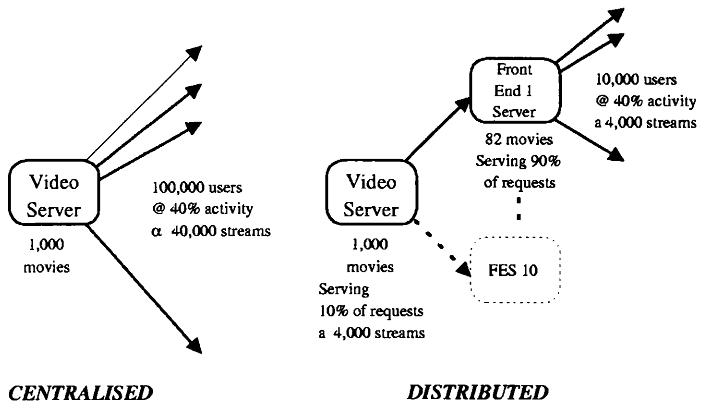

Figure 5.9: Centralised versus distributed system

approach is actually cheaper in terms of storage than the centralised approach for a total of 40,000 active video streams. It is true, however, that the larger the centralised server becomes, the more cost effective storage becomes.

The results of course are valid under the particular assumptions used in this study, especially on calculating disk cost. But the authors conclude that the results expressed above, however, are sufficient to conclude that a distributed approach to VoD is preferable for a number of reasons and costs, such as in terms of storage and bandwidth usage.

In this paper the distributed storage solution is proved to be more efficient than the centralised one. Although storage cost is difficult to predict, the savings in network bandwidth and usage justify the conclusion. The most efficient distributed system is when 85% of the requests is serviced locally. They use the classification of movies according to popularity and show that the bulk of requests is destined to a small number of movies, a point to which I shall refer later. They propose the distributed system as the most cost effective one as is done in this study.

5.3.4 An Architecture for interactive applications [4]

In this work, the three main components of the proposed architecture are: the broadband Backbone Network, the Wideband Access Network and the Public

Switched Telephone Network. In order to support the inter-working of the broadband backbone and the slower local access capability at the periphery of the "inner" Broadband network, they propose the deployment of network based facilities which they call Service Circuits. The Service Circuits can also serve as a platform upon which a portion of the application is run. Later, with the deployment of broadband access these applications will migrate to the user terminal equipment.

In the proposed architecture the Service Circuits are not integrated with the switch fabric and therefore may be located either inside or outside the local exchange, see *Figure 5.10.*

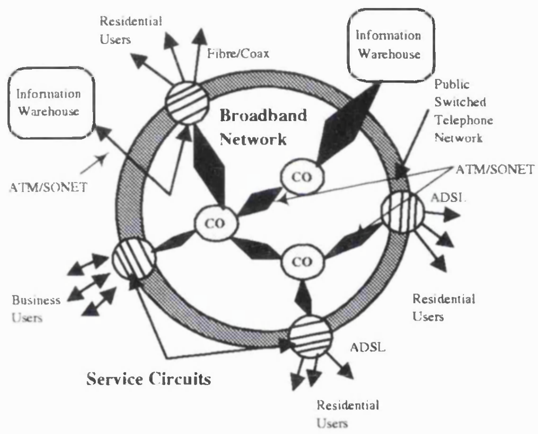

Figure 5.10: Overall Service Network Architecture

VoD is the application examined in more detail. Video and audio information is compressed according to the MPEG standard, and is stored in the Information Warehouses (IWH). Magnetic tape is used for archival storage and magnetic disk arrays for on-line storage. Information is transported from the IWH on demand in fast bursts via the Inter-Office high speed links to the local exchanges where the original temporal properties of the information are stored. It is then delivered to the customers over the access line in real time.

The Service Circuits located at the local exchanges act as adapters for the entire protocol stack between the BISDN and the Wideband access networks.

They consist of per-customer Buffer Line Cards which single board computers optimised for real-time applications and for which software is downloaded from the network.

The actual information retrieval is managed by the Buffer Line Cards which place requests for bursts of data from the IWH via the supervising local exchange processor. The Line Cards buffer the ATM/SONET transported data for subsequent playout to the user.

When a user selects the VoD application the processor that runs the Buffer Line Cards places the user's request with an IWH. If the request is granted, the IWH downloads an appropriate Media Presentation Script into the Buffer Line Card. The local exchange service processor orders the start of the application script execution.

The part of the application that runs in the user premises equipment is responsible for the user control interface, for the MPEG system stream parsing and the part running in the local exchange is the VCR emulation software. The application portion running in the IWH is responsible for issuing the application script, and upon selection of the material, issuing the material map, and then for handling data requests received from active users.

The proposed architecture supports the idea of delivery distribution to the local exchange level, but is not completely decentralised as part of the service depends on the IWH. Distributing the IWHs and providing more functionality to the local exchange service processor would make service provision more flexible. The use a broadband ATM network, is possible using the appropriate protocol such as of the one proposed in the paper. The introduction of service processing distribution and the use of a broadband network, make this paper one of the earliest in proposing a distributed VoD structure.

5.3.5 Summary of reviewed methods

In the works reviewed here, a distributed system to model VoD, has been selected. All papers demonstrate the efficiency of distribution. The service, and the anticipated use of the service, as well as the size of current networks, dictate the use of a distributed solution. A network of interconnected servers, provides more flexibility to the operator and higher availability to the end user as demonstrated in section 5.3.1 on p. 115 and in 5.3.2 on p. 118. By placing the requested material as close to the end user as possible we make multiple savings, in bandwidth usage, delays due to switching speeds and improve on service reliability, since the user depends on his local server only. Additionally, the distributed solution is not more costly than the centralised one see section 5.3.3 on p. 123, as well as offering the above presented advantages. In future broadband networks, should support multimedia servers placed on locations that impose minimum load on the network infrastructure as argued in one of the first papers see section 5.3.4 on p. 125.

In our work we take this argument further and provide a design that not only distributes the material but the control as well, in order to exploit further the features of the dynamic VoD system.

5.4 Summary

In this chapter we have analysed the most important features of a VoD distributed server network. For the service to work efficiently and provide satisfactory QoS to the end users the majority of service calls should be answered successfully. But all calls are not made out for the same movie. Most calls require a small number of movies, and the rest are divided among a large variety of choices. For the service to work satisfactorily, most calls should be answered, even those for less popular movies. This though should be done considering network usage efficiency. The most important resources in the VoD service are the VoD servers and network bandwidth. Both should be used as efficiently and economically as possible to make the service a viable economic proposition. In the next chapter we analyse our proposed control protocol for such a system based on the service and design principles analysed in this chapter.

In the next chapter we shall present our management scheme for a distributed network of the above VoD servers, based on network resource utilisation.

Chapter 6

Management Scheme for distributed IMS

6.1 Introduction

Having analysed the advantages and disadvantages of centralised and distributed systems in the previous chapter, in this chapter we analyse our proposed management system for a distributed VoD service. The system is applied on simple square grid network topology.

6.2 Distributed VoD and management system

As already stated, our design for VoD is focused on a system that can be applied commercially, and to a broad customer base. VoD belongs to the newly established family of multimedia services. Although a strict definition of multimedia does not yet exist, Wolf *et al.* present the following definition in [57]: "Multimedia denotes the integrated manipulation of at least some information represented as continuous media data (such as video and audio data), as well as some information encoded as discrete media data (such as text and graphics). The "manipulation" refers to the act of capturing, processing communicating, presenting and/or storing".

The VoD services makes use of both types of media, with the use of selection menus and the use of films as the offered material. Although in our system we are primarily concerned with the delivery of continuous media data, the system can easily expand to incorporate the delivery of discrete media data.

The underlying feature of VoD, is the process of interactive data retrieval. VoD is only one IMS service. As such it shares some generic features with the rest of the services belonging to the same family. The most important characteristics of such a VoD service are summarised as follows and in *Figure 6.1*:

- Simultaneous and synchronous delivery of image and voice to the end user.
- Real-time control over the delivered material available to the end user.
- Multiple access requests for the same material. Usually these requests are done asynchronously and during a certain period of time called "peak time".

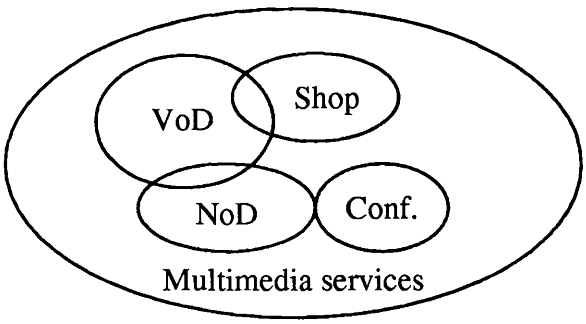

Figure 6.1: Common features of multimedia services.

The core bandwidth requirements for delivering the multiple data streams to the users are immense and user number dependent. Legacy systems are not built for this environment and will not be able to handle the transition. The typical duration of a movie is 100 minutes [58] compared to the typical phone call duration a few years ago which was a few minutes, makes the need for bandwidth dimensioning obvious.

The user–data access relations will be quite complex. Demand for each piece of data will be different. In the case of movies this relation will be "success" and time of day depended as market research has shown [59]. For simplicity of description we have simplified the continuous range of popularity to three major classes. These we shall use in our simulations.

- **The very popular movies.** These are typically blockbusters at the peak of their demand. This is the material most people would want to watch on the big screen.
- **The popular movies.** These will attract a smaller number of users. They will be ex–very popular movies for which demand is waning or classic movies.
- **The "specialised" movies.** These will attract only a small number of requests and could be priced accordingly. Specialised movies will be essential for the service to provide, as they could constitute a sizeable proportion of the service income. From the service designer point of view, they provide a model for other IMS services such as an interactive library.

The advantages offered by distribution make it suitable for the heterogeneous environment of multimedia services. These are:

- Connection placement as close to the end user as possible. The servers taking advantage of any localised effects in movie popularity as well as reducing bandwidth utilisation.
- Flexibility and robustness. In the event of failure, the users can be routed to another server while their local server is being repaired.
- Easy expansion and upgrade. Adding new servers to the system will increase its capacity without the costly implications of replacement. Upgrade can be done gradually and the cost is spread over a period of time.
- Heterogeneity of servers. Servers of different capacity may be used to cater for the different types of information and provide the end user with different QoS and functionality [60].

To use the advantages offered by distribution we have chosen to model the VoD service based on a distributed VoD server network.

Due to the size of the system all state information is not useful to a particular server except in some specific cases as we shall see later on. To support a good QoS a small service blocking probability is required. This amounts primarily to the users accessing the very popular material. These users will render high profit to the service. On the other hand there exists the seldom accessed material, which can produce profit if handled appropriately. Such an infrastructure, would be flexible enough to support other IMS services too. This type of material could be accommodated too using the appropriate control method. This is what we have tried to achieve in our proposed algorithm. A system that can respond to a range of access patterns, from very popular material to least popular material while managing its network resources as efficiently as possible.

6.3 Management system

The control or management system is an integral part of the distributed server system. The proposed protocol should be able to cope with the three different types of movies and distribute them accordingly to the servers. Multiple copies are created according to user demand. Remote connections should be established if necessary. Server capacity should be managed in such a way that the copying process does not exceed the minimum required. The robustness and the flexibility of the system will be determined by the robustness and flexibility of the control protocol used. The control tries not to overload network bandwidth by introducing extra signalling, so that the protocol is efficient and robust. The amount of information required in decision making is kept at a minimum, to make decision making simple in processing requirements. Network resource management is to be performed by the management system.

Due to the complexity of the system we have chosen to study a simple regular square grid network. Any real world network irrespective of complexity

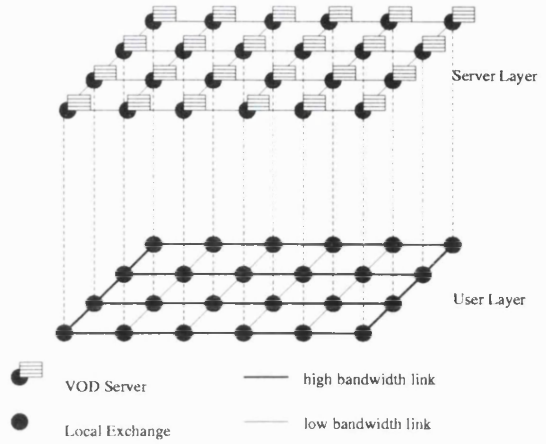

Figure 6.2: User network and server network connection.

can be mapped into a logical square grid network. We used the most general of all topologies in order to apply our proposed management system. In this way we can examine the most important control parameters of a distributed server system. In our simple model each server is associated with a local exchange *Figure 6.2*. The local exchanges are connected via high bandwidth links whereas the VoD servers are connected via low bandwidth links and no image traffic is routed between them. We use this network as a platform for the construction of a control protocol needed to optimise the use of network resources.

We are using this structure to assist in the construction of a control protocol needed to optimise the use of resources. In order to test and evaluate our control procedures, we have outlined and built a basic system of networked servers. The square grid network is an abstract example which can be mapped to any real world network.

6.3.1 Principles

In building the control protocol the basic principles had to be analysed. The following principles have been used to construct, modify and test the proposed control protocol:

1. Multiplicity of offered material. In order to keep bandwidth consumption at low levels multiple copies of the same movie are desirable. Bandwidth usage is directly proportional to the distance between user and serving node in hops. A simple calculation using random creation of copies shows how the average bandwidth usage relates to the number of extra copies available in a 3×3 square grid network, (see *Figure 6.3*). Random positioning though is not a desirable feature in a system aiming to exploit user demand distribution. The multiplication process has been modified so as to follow user demand, the new copies are created on the nodes with the highest demand for them. As the VoD server is specified to be able to play more than one movie simultaneously, the case of replication on the same node had also to be considered. This strategy though is not favoured, since it greatly reduces the server's ability to satisfy diversified demand.

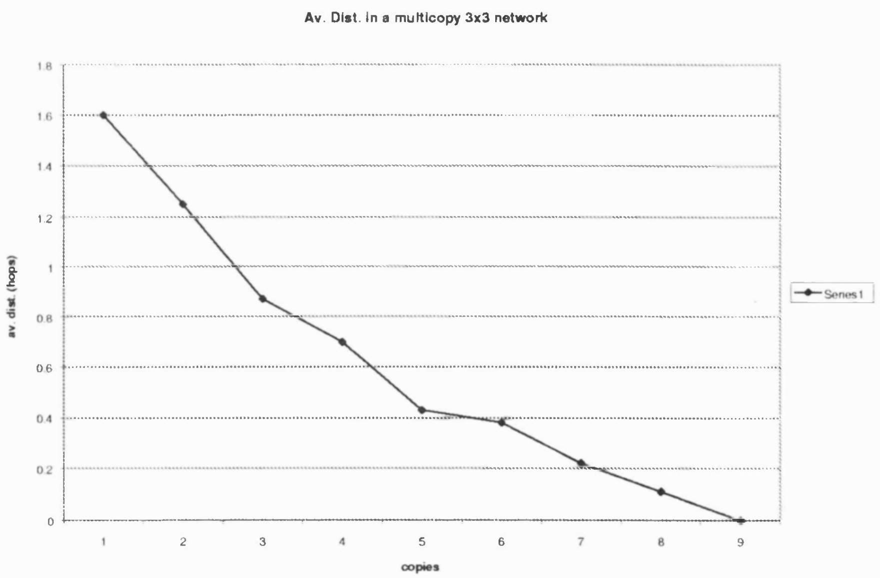

Figure 6.3: Relation of average distance to the nearest copy as the number of copies increases.

2. The minimisation of bandwidth usage. With the introduction of optical fibre technology, bandwidth may not be such a scarce resource in the

future. But as new IMS applications are introduced and their popularity increases, they will fill the available resources. Generally it is likely that bandwidth will always be a precious resource and should be used as efficiently as possible.

3. The usage of the VoD server capacity. The system's ability to cope with diverse user demand depends on the server capacity and the protocol for its usage. The total VoD server capacity determines the diversity of the offered material. In the proposed architecture we use a system of networked VoD servers, therefore the protocol used will define the number of different movies simultaneously delivered. As a first approach to this direction a keep–local protocol is used. Before the loading of a new movie on a sever, the servers in a predefined area are checked for possibility of connection. If the appropriate material is found there, then that connection is favoured to the local one. The dimensions of the search area relate to the extra bandwidth used. The extensions and limitations of this strategy are currently under study.

4. Quality-of-Service (QoS). An important factor for the service viability and success is the QoS offered to the end user. The role of QoS should not be undermined since it is vital for the acceptability, success and competitiveness of a commercial system. The VoD service will have to compete with other existing forms of movie watching, such as the broadcast and Cable TV and the local video club. Failure to connect, abrupt disconnection while watching a movie and deteriorating image quality while viewing are some of the factors determining the QoS offered to the end user.

Once these parameters and principles have been identified, we have a testbed for further simulation and analysis.

6.4 Summary

In this chapter we have presented our distributed management system and the principles it is based upon. These are meeting user demand under a negotiated QoS value, while using as effeiciently as possible the two most important system resources. In the next chapters we shall apply this system on a hypothetical VoD server network, under varying user demand and server size we shall examine and compare the results against a centralised system.

Chapter 7

Simulation framework

7.1 Introduction

In this chapter the management system simulation framework is described. We start by giving the logical flowchart on which the management system is based. We then give the flow of logical decisions needed for the system to function. We explain the decision making process the system uses to manage service operation and network resources. The most important service parameters are identified and evaluated with respect to service operation and network resources usage. These parameters form the basis of the management system. The cost function presented serves as the comparison criterion on how these parameters influence system operation.

The management system operates using network information, which is then translated into actions through its own decision making process. Its operation depends on the minimum possible pre-stored information on network and server status, most of the required information for its operation is acquired on-line and on demand. The management system is based on a search procedure to establish user connection to the required movie and a relocation process, which is evoked to improve system operation. These procedures are explained in the simulation framework description along with the conditions and restrictions on their use.

7.2 Algorithm Description

As argued in the previous chapter, a distributed management algorithm will be essential for the implementation of a VoD service. A simple calculation of the centralised bandwidth shows the need for the implementation of a distributed approach, if not from the beginning of service implementation, then very soon after it has been established in the market. In the European Union alone, during peak time there are approximately 200,000,000 viewers in 100,000,000 viewing households for both terrestrial and cable television. Using the data rates of 2 Mbit/s for MPEG-2 and 6 Mbit/s for HDTV transmission and assuming only one independent stream per household, the total aggregate bandwidth required during peak viewing periods is 200 Tbit/s and 600 Tbit/s, respectively. This requirement by far exceeds the available backbone bandwidth offered by both telecommunications operators and CaTV providers. This huge amount of bandwidth clearly dictates the need for efficient bandwidth management. In the previous chapter, we introduced the principles upon which a VoD management system should be based, in this chapter we analyse how this system works and how we use it to extract information on service operation and QoS provided to the end user.

7.2.1 Simulation framework

We use the structure described previously, a network of VoD servers each associated with one local exchange. Although, this is the simplest example, it can be easily expanded and adapted to a more complex scenario. This level of distribution is not essential and binding, but is used as an indication of where in the network hierarchy the video servers could be placed. The actual location of the servers depends on network infrastructure. A hierarchical network structure can be used, consisting of a national backbone ATM network, regional or metropolitan ATM networks, and local loops, which are connected to the regional or metropolitan networks, via head end switches, as shown in *Fig-*

ure 7.1. The number of STM-16 backbone trunks and STM-4 regional trunks that would be required for an example scenario in support of MPEG-2 transmission depends on the number of users that connect to the head end servers. Archival video servers may be attached at any point in the hierarchy. Video distribution is accomplished on a Multipoint connection tree, imposed on the MAN.

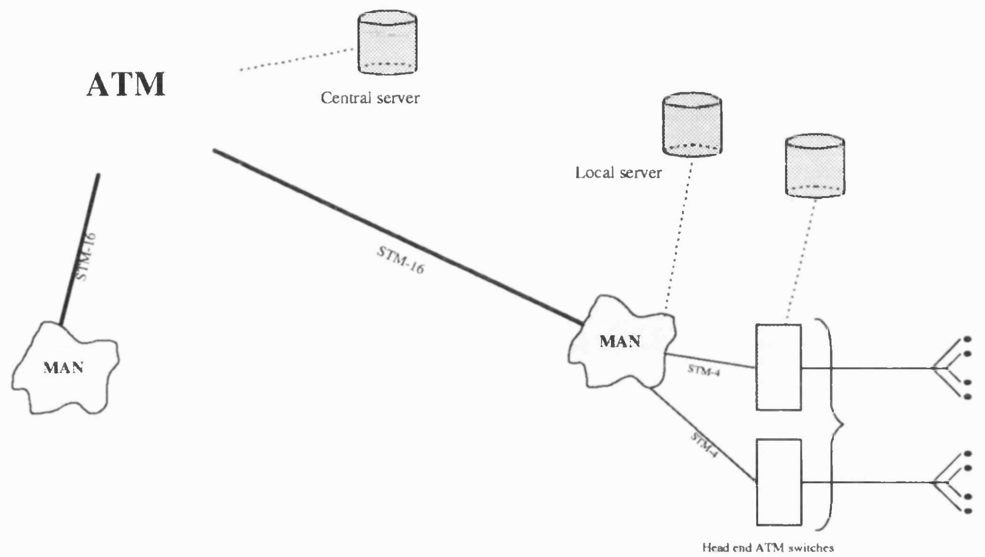

Figure 7.1: Network model.

In the current CATV tree broadcast infrastructure, the number of households connected to a single tree has a significant effect on the Quality of Service provided to the end user. If there are too many users connected to a single tree, many users may be denied access to interactive multimedia. Therefore, the number of households connected to a single broadcast CATV media offering interactive services must be limited to fewer than 1000 in the case of a 2Gbit/s tree. This will continue to be true in future ATM switched network deployments in order to limit the head end switch to a reasonable size (cost), such as 1 K $\times$ 1 K. Thus, the distribution tree is limited in breadth.

The control algorithm operates on top of the server network. The control algorithm is used to manage network resources, to achieve the QoS requirement. In a distributed VoD server network there are two important resources to consider:

- **Core bandwidth.** Although core bandwidth availability is bound to increase in the future, bandwidth will never be an abundant resource. Efficient bandwidth management will provide larger profit margins to service providers.

- **Server load.** The VoD server is the other element of the service, which although essential, will not be cheaply available, especially at the beginning of deployment. Efficient server usage will allow the service provider to offer good a QoS, and maintain low operation costs.

As argued in the previous chapter, placing the servers close to the end user reduces core bandwidth requirements and localises traffic in the periphery of the network.

The control algorithm loads and unloads the movies onto the servers and establishes local and remote connections. Each server has a certain capacity C_s, which is usually a fraction of the total number of movies held in the library C_l. The total server system capacity C_{ts} is varied with respect to the number of movies in the library, this is done so we can study the process of distribution, and creation of multiple copies. Usually local demand cannot be satisfied by the local server so remote connections have to be established by the management protocol. The management protocol is based on five procedures:

- Network search for available movies, during which connections are established to movies already loaded on local or remote servers.

- Network search for available server space, during which new movies are started on local or remote servers.

- Movie removing from server, during which a movie which is viewed by a small number of users is labelled as "going-down" and no new users can attach to it. The movie is removed when all attached users have finished watching. If a sudden change in movie popularity occurs, the movie is reinstated in playing status and new connections are allowed.

- Movie relocation, during which not very popular movies are relocated to a distant server in order to free server space for a more popular movie.
- Movie discontinuance that is only evoked under extreme server space scarcity; a movie is stopped to make space for a more popular one. This happens only if the relocation and "going-down" proecesses fail. The users are abruptly disconnected due to this network management decision.

7.2.2 Network search

Each server receives user requests for various movies from its local user pool. If the server is able to provide the movie locally from its own set of playing movies, the users are connected locally. If the server cannot provide the requested movie, then depending on user demand, a cluster of servers is searched and if the movie is found available, the users are connected to the appropriate server. The radius of the search is inversely proportional to user demand. This limits high bandwidth usage to a small area, and allows for better server space management since the least popular movies are not repeated in more than one server. The search radius is varied from one hop to three hops for the popular movies, whereas for the least popular movies a search of the entire server network is allowed. The benefits of the search range dependency on popularity are demonstrated in [61].

In the case that the requested movie has not been found playing on the allowed server cluster, the local server is checked for available server space to start the movie locally. If no free server space is available then the server tries to start the movie in another server located in a predefined server area. This area is the same as used in the retrieval search. The movie is started on the first server encountered with free server space and the users are connected there.

If this search procedure fails, the users are not granted their request, unless relocation is used.

7.2.3 Relocation

If a movie has not been found within the defined search range or no free space is available for it to start, we have introduced a relocation scheme. Relocation is used to make a small-scale redistribution of the available movies, in order to provide free space for a new movie to play. The server selects the movie with the least attached users and performs a full network search to find another instance of this movie. If the movie is found playing on any other server, the users attached to it are relocated to that distant server and free server space is created in the local server, so that the new movie can start there. This relocation method is used for the least popular movies in order to keep network bandwidth usage to a minimum. This reduces the number of copies of the less popular movies, increases the number of copies of the more popular movies and provides more efficient network resource management as will be shown later.

The relocation method though requires extra network bandwidth during the location transition. The extra bandwidth should be available to the system in order to support the smooth transition from one server to another and preserve the quality of the picture delivered to the end user. To further reduce the problem, extra frames could be stored at the CPE (Customer Premises Equipment) buffer and be played during the transition time to avoid deterioration in image quality. In the simulation we have not addressed the problem, since appropriate underlying system architecture can solve it.

7.2.4 Movie removal

To anticipate falling movie demand we replace a number of "unpopular" movies with other more popular ones, movies with a small number of attached users are slowly removed from the system.

In the beginning of each time step, the server checks which movies fail to attract fewer users than a set threshold. These movies are labelled as "going-down" movies. The procedure for taking down a movie has to ensure that the QoS is preserved during the process. In the "going-down" movies no new connections are allowed and after a number of time steps, which is the same as the number of segments the movie is divided into, in our simulation, the movie is removed from the server, making space for a new movie to start. The movie is removed when the already attached viewers finish watching. To anticipate a sudden change in user demand, the "going-down" status is removed if the number of users from one local exchange which want to attach to the movie at a given time step exceed a set threshold.

7.2.5 Forced movie termination

If the relocation process fails to connect the users anywhere in the server network, an extreme measure is taken. The local server selects a movie, which is cut abruptly to make space for the new movie to start. The movie selected, is a movie that has been playing for more than a certain time limit and the users attached to it fall below a certain threshold. The time limit criteria is used to avoid a fast movie changing process and allow time for the movie to gain popularity, as for example at the beginning of the afternoon-evening peak time. The user limit is used to avoid disappointing too many customers, as this system management action will result in very disappointed users and consequently deteriorated QoS.

The management system operation, using relocation is summarised in the flow chart shown *Figure 7.2.*

7.3 Pseudocode

The management system action is summarised in the following pseudocode, for the management scheme with and without relocation.

This is the pseudocode for the management system with the relocation pro-

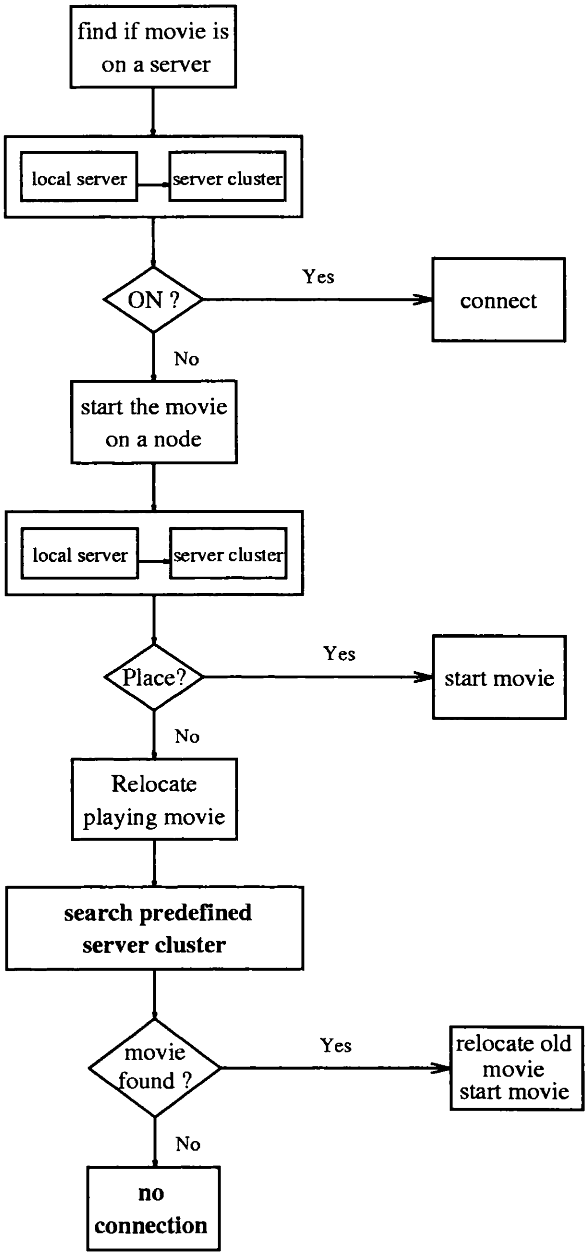

Figure 7.2: Simplified flow-chart.

cedure unavailable.

```
for all servers all movies perform control action;
 connect users to a movie already playing;
  either
```

```
    on the local server
     lift the ``going-down'' flag if demand is enough;
   or
    on locally placed server;
  if action fails
  connect users to a movie now staring;
   either
    on the local server;
   or
    on locally placed server;
  if action fails
  no connection possible;
   return users to user pool;
```

This is the pseudocode for the management system with the relocation procedure available.

```
for all servers all movies perform control action;
 connect users to a movie already playing;
  either
   on the local server
    lift the ``going-down'' flag if demand is enough;
  or
   on locally placed server;
 if action fails
 connect users to a movie now staring;
  either
   on the local server;
  or
```

```
   on locally placed server;
 if action fails
 relocate the movie with the least users attached;
  start the new movie in its place;
 if action fails
  take down a movie with attached users below threshold;
  start new movie in its place;
 if action fails
 no connection possible;
  return users to user pool;
```

7.4 QoS factor

As VoD is a customer-oriented service, customer satisfaction is an important factor for service take-up and consequently service turn over. QoS is the term usually used by service providers to determine user satisfaction with the provided service. In a multimedia service such as VoD, QoS is used as a term indicating all those aspects of service provision, which the user considers essential to prefer this particular service from another.

The term QoS includes quality of image delivered to the set top box, user connection success and availability of interactive functions. As a consequence of using a management system such as the one proposed here, a number of user requests may not be satisfied due to the control system management actions.

A metric based on the number of failed connections is used, as a partial indicator of the provided QoS. As described previously, the control system may even disconnect users under certain conditions. This action although useful for network resource management, results in extremely dissatisfied users. Although the provider may try to alleviate the bad impression by offering counter-benefits such price discounts, the fact remains that abrupt user disconnection constitutes a major factor in deteriorated QoS.

We express this particular aspect of QoS resulting from the implementation of our management scheme, as the normalised percentage of the failed and cut connections. The cut connections carry more weight and are multiplied by an importance factor determined by the service contract negotiated between the service provider and the customer.

7.4.1 Comparison criterion–Cost function

In order to assess the performance of the management scheme under changing parameters, we define cost functions for network bandwidth and server cost. An overall cost function can then be computed, which will be taken as the comparison criterion. The partial cost functions are normalised so that different parameters can be computed and analysed.

Furthermore, we define a cost factor to compute non-linear pricing of bandwidth and servers. Finally, the cost function is weighted to evaluate different cost ratios between storage and bandwidth. The storage cost incorporates the cache server and the necessary storage costs.

1)*Bandwidth Cost*: Bandwidth on the links that connect the servers is proportional to the number of movies playing on the server and the number of users attached to it. The total bandwidth (C_{tb}) is the sum of bandwidth of the individual links over the number of links in the network.

$$C_{tb} = \gamma_b \sum_{\Lambda} b(\lambda) \tag{7.1}$$

where Λ is the set of links, $b(\lambda)$ is the bandwidth on link λ, and γ_b is a weight factor. Since C_{tb} is normalised, the cost of bandwidth to serve all the users at a head end can be arbitrarily chosen, i.e. may account for a real cost value when available.

2)*Server Cost*: The video server cost is dominated by the number of movies it makes available to the users. Efficient utilisation of server space is an important parameter in the running of the service. Therefore, the server cost function can be expressed as the normalised percentage of server capacity utilisation.

$$C_{ts} = \gamma_s \sum_S z(i) \tag{7.2}$$

where S is the number of servers in the network, γ_s is a weight factor, $z(i)$ is the cost of a server with capacity i movies.

3) *Total Cost*: The total cost of the system is the sum of the total costs of storage and bandwidth

$$C_t = C_{tb} + C_{ts} \tag{7.3}$$

In order to evaluate the dependency on the non-linear costs of links and servers, we define a cost factor $\phi = (\phi_s, \phi_b)$, where ϕ_s is the factor for the storage components and ϕ_b is for the bandwidth. We assume that in terms of link cost, there is a saving as higher capacity is deployed, i.e., the price per unit of bandwidth is lower as the capacity of the link increases. In this case

$$C_{tb}(\phi_b) = \gamma_b \sum_\Lambda b(\lambda)^{1/\phi_b} \tag{7.4}$$

where $\phi_b \geq 1$. At the same time we assume that the server cost has the inverse property: it is even more costly per unit when higher demands are to be supported due to extra running and construction cost ($\phi_s \geq 1$). Therefore

$$C_s(\phi_s) = \sum_S z(i)^{\phi_s} \tag{7.5}$$

Using this cost factor, the total cost then becomes

$$C_t(\phi) = C_{tb}(\phi_b) + C_{ts}(\phi_s) \tag{7.6}$$

Furthermore, we consider different cost ratios between the total bandwidth cost and the total server cost in order to evaluate their relative impact. Since we measure individually bandwidth and total server cost, we can easily capture this difference by means of a weight factor in which case, the overall cost in this case is given by

$$C_t(\rho) = \frac{2}{1+\rho}\,(\rho\, C_{tb} + C_{ts}) \tag{7.7}$$

where $\rho > 0$ provides the relative cost ratio between bandwidth cost and server cost. For $\rho > 1$ bandwidth becomes more important and for $\rho < 1$ server cost dominates.

In the cost function graphs presented in the next chapter, both bandwidth and server costs are of equal importance. Therefore, $\phi_s = 1/\phi_b = 1.1$ and $\rho = 1$. The bandwidth and server factor have been chosen to provide a small non-linear deviation.

This cost function is used for our server network, under varying server capacity, changing user demand and different management system parameters. Using the cost function we can extract results for the optimum network and management system parameters to help us correctly dimension our network, to honnor a QoS agreement. QoS can be expressed as the sum of the normalised successful user connection attempts and the normalised number of abruptly disconnected users.

$$C_{qos} = C_{suc} + C_{cut} \tag{7.8}$$

As the second term carries more weight in user perception of a reliable service we can introduce a weight factor ϕ_c.

$$C_{qos} = C_{suc} + \phi_c C_{cut} \tag{7.9}$$

where $\phi_c < 0$ since the influnce of the cut users will be negative to the ocerall QoS value. In the QoS graphs presented in the next chapter the value of the cut user weight factor used is $\phi_c = -1000$, since the number of cut users is very important in shaping a negative QoS value.

The proposed cost function with the appropriate parameters correctly set will be used as a measure of the total cost of service implementation. The importance of the cost function parameters will depend on the actual network and server costs and will change according to the particular situation. Furthermore, the cost of obtaining the offered material and the copyright need to be taken into account.

7.5 Summary

In this chapter we analysed the simulation framework, we used to run our algorithm. We described the management actions of the control algorithm. These include the starting of a movie in a server, local and remote connection establishment and movie relocation. Movie relocation is used as a simple method to improve resource utilisation. Local and remote connections are established depending on movie demand and server load. The management algorithm performs a different connection procedure depending on the number of users requesting connection to a particular movie. Relocation is used for a special category of movies and provides a mechanism for flexible server space management.

In the next chapter we present and evaluate the performance of the proposed algorithm. We vary the most important parameters, such as search distance range and server capacity. Then we analyse and compare system performance when no relocation, limited and full relocation management systems are used.

Chapter 8

Performance of the proposed algorithm

8.1 Introduction

Using the simulation framework described in the previous chapter, we assess the performance of the proposed VoD service. We vary server capacity and server search range and apply our proposed management systems. We present our results and compare management system performance with and without the relocation process. We monitor and present average bandwidth usage, server capacity usage and the QoS provided. QoS is based on successful connection and abrupt disconnection rates.

Using the results derived from our simulations, we can dimension network bandwidth and server capacity in order to provide a VoD service meeting a preset QoS requirement. In all cases we have used a changing probability profile for the movies stored in the movie library. This profile has two time varying states and user demand oscillates between the two. This change simulates periodically changing user demand that might occur during a 24-hour period.

8.2 Simulation

8.2.1 Probability generator

The size of the movie library was chosen to be 100, as a number that provides a good variety of material to the customer without imposing a huge server space demand on the network. The movie popularity function used, was a three step function dividing movies into three categories according to popularity. The very popular, the popular and the least popular. In the very popular movies the high street cinema movies are included, in the popular movies older movies and less popular movies are included and in the least popular specialised, old and movies addressing a very specialised audience are placed.

In most literature in the area, the probability profile is given by a function based on Zipf's law [62, 63, 2]. Zipf's law is defined as: given N_p programmes, ordered by its popularity, the probability that the ith movie is selected is given by $z(i) = C/i$ where $C = 1/\sum_1^{N_p}(1/i)$.

But as these are the early days of studying traffic generation between offered material and user demand, not everyone agrees to the accuracy of the representation of user demand given by Zipf's law [3]. We chose a stepped function, which provides a clearer distinction among the different movie categories, from the high street material to the specialised movies. *Figure 8.1* shows the cumulative distribution function for the Zipf and step function, where this difference becomes clear.

The total service access probability for the assigned customer base is 93.85%. The movies in the library were assigned a different popularity value according to their class. The very popular ones were assigned 0.1% the total probability for the 25 movies in total adding to 92,82%, the next 25 popular movies were assigned a probability of 0.0048% adding to 0,80% and the last 50 least popular movies were assigned a probability of 0.0008% adding to 0,23% of the customer base. The total probability was calculated by sequentially applying the probability profile to the customer base. The total service access probability was high

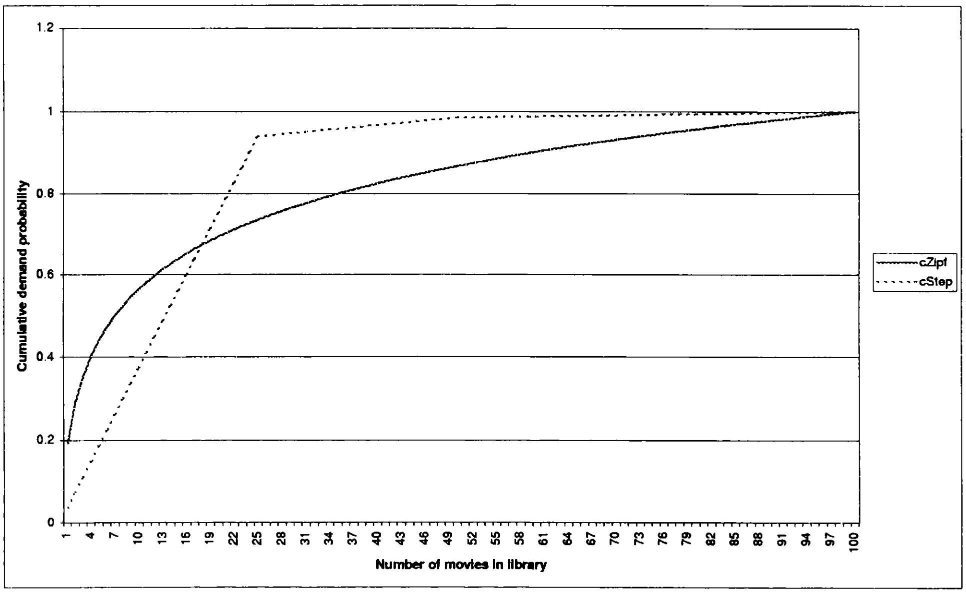

Figure 8.1: Cumulative Distribution graph of Zipf and stepped functions

and this forces the system to work under full load conditions. Additionally, the stepped probability profile becomes completely inverted for a third of the simulation time, while keeping the cumulative probability constant and gradually changing demand. The most popular and popular movies lose demand until they become least popular and the least popular gain demand until they become popular and most popular, the transition takes 40 time steps to complete. This inversion of the probability profile is introduced for two reasons; firstly to simulate changing demand during the 24-hours of the day and secondly to monitor system performance and give the provider extra information on how to dimension the network.

8.2.2 Simulation parameters

The simulation was run varying server capacity, movie search distance and relocation search distance. Server capacity was varied from 1 to 15 movies simultaneously available per server keeping the size of the library constant at

100 movies. Server size is one of the most important parameters in calculating service cost; increasing server capacity improves movie availability but adds to the total service cost. Keeping server size as a parameter we monitor how server size improves service quality.

The movie search distance was varied from 1 to 3 hops, for all but the least popular movies, where a different search distance was applied. Searching for an existing copy of a movie or even starting it within the area of a server cluster, improves server space utilisation and should keep bandwidth usage more uniformly distributed throughout the network, since the servers in the cluster, can cooperate in order to satisfy local demand. Extensive searching may introduce extra bandwidth usage due to the number of remote connections established. Varying the search distance, we monitor service performance against bandwidth usage. For the least popular movies full network search is allowed as core network bandwidth usage will be small. Relocation is used as a simple management method for the network to improve on movie and connection distribution. When a movie is relocated, a number of users change VoD server, and are connected to another remote server. Relocation can improve movie availability, but is expensive in terms of network resources as a number of distant connections are created. We introduce the relocation process and vary its range from half to full network radius. We monitor the effects of relocation on network performance and QoS offered. The results obtained are used to evaluate its contribution to improved service performance.

8.3 Results

In the results presented here we analyse how the most important system parameters behave under different control algorithms.

The system parameters we have identified and used their values to measure control system performance are the following:

- bandwidth usage, here defined as the number of hops occupied per connection established, used as an indication of the traffic load imposed on the network due to VoD service operation,

- the number of user requests the control system failed to complete as an indicator of successful service provision to a customer base and as a measure of the QoS offered to the end user,

- the number of users abruptly disconnected due to a management system decision as a measure of the QoS offered to the customer base, since abrupt service disconnection is considered to deteriorate service performance heavily

We use a 10×10 server network with 100,000 users per server location, the capacity of the local server is varied from 1 to 15 movies and the number of segments the movie is divided into is 5. Although, the number of segments chosen was small and more appropriate for a n-VoD service, it was used here as a representative number since we do not model the load imposed on the network due to interactive user actions. The system is monitored for 120 time steps. The performance of the system was evaluated without the relocation process and with relocation using a search range varying from half the maximum distance of the network to full network search.

First, as shown in [61] we demonstrated how full network search for the least popular movies improves connection success rate.

System performance with increasing search distance and constant server capacity

Next we examine the effects of increasing the search range from 1 to 3 hops, for the non-relocating, half-relocating and full-relocating scenario keeping constant the local server capacity.

Using 1-movie server capacity Here we examine the effects of increasing the local search range from 1 to 3 hops, on a one-movie server capacity network, for the non-relocating, half-relocating and full-relocating scenario.

In the 1-movie server capacity case, bandwidth usage increases as search range increases from 1 to 2 hops, see *figure 8.2*. In the full-relocating system the increase is less steep than in the other cases, of the order of 20%, whereas in the half-relocating case the increase is of the order of 40% and in the non-relocating system the increase is steeper of the order of 50%. As the search range increases even further from 2 hops to 3, the increase in bandwidth is steeper; in all three cases is of the same rate of 40%.

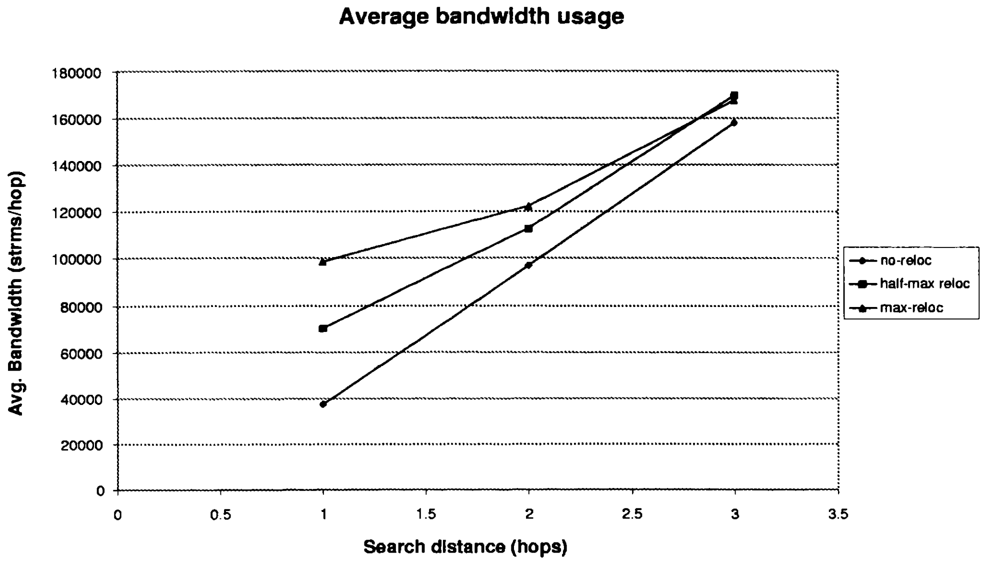

Figure 8.2: Average bandwidth usage using severs of capacity 1

The connection success rate, increases as the search range increases from 1 to 2 hops, see *figure 8.3*. In the non-relocating case the connection success rate is higher and increases by 30%. In the half-relocating system, with lower success rate the increase is of the same order as before of 30%. In the full-relocating system, where the success rate is the lowest, the increase rate is lower of the order of 15%.

Finally, as the search range increases from 2 to 3 hops, the connection success rate continues to increase, see *figure 8.3*. In the non-relocating system, with the highest rate the increase is of the order of 10%. In the half-relocating system the increase is of the order of 20% and in the full-relocating system, with the lowest connection success rate the increase is of the order of 25%.

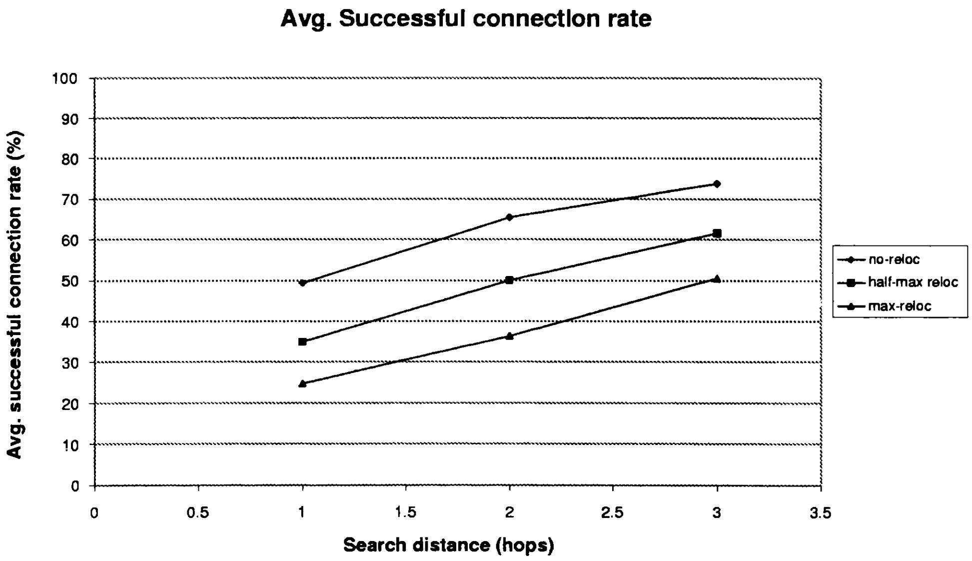

Figure 8.3: Average success connection rate using servers of capacity 1

The cut user percentage is almost zero for the half and non-relocating systems, see *figure 8.4*. In relocating cases the cut user rate is very small 1.2×10^{-7} for the 2-hop search distance and 4×10^{-8} for the 3-hop search distance.

The cost of the system increases as the search range increases from 1 to 2 hops, see *figure 8.5*. In the full-relocating system, where the cost is the lowest, the increase is of the range of 30%. In the half-relocating system, where the cost is higher than before, the increase is steeper of the order of 40%. In the non-relocating system, where system cost is the highest, the increase is of the order of 45%.

As the search range increases even further, the increase in system cost is of the same rate of 20% in all cases, with the full-relocating system being the least expensive.

The QoS offered by the system follows the same trends as the successful connection rate, see *figure 8.6*.

As the system is under-dimensioned, the best performance is seen in the non-relocating network. Only the more popular movies find a place on the server

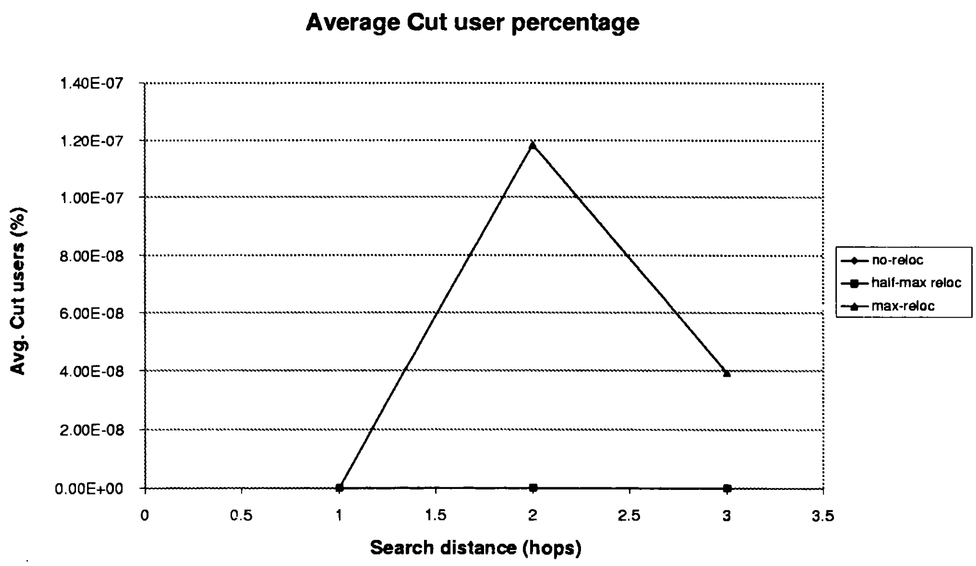

Figure 8.4: Average cut user percentage using servers of capacity 1

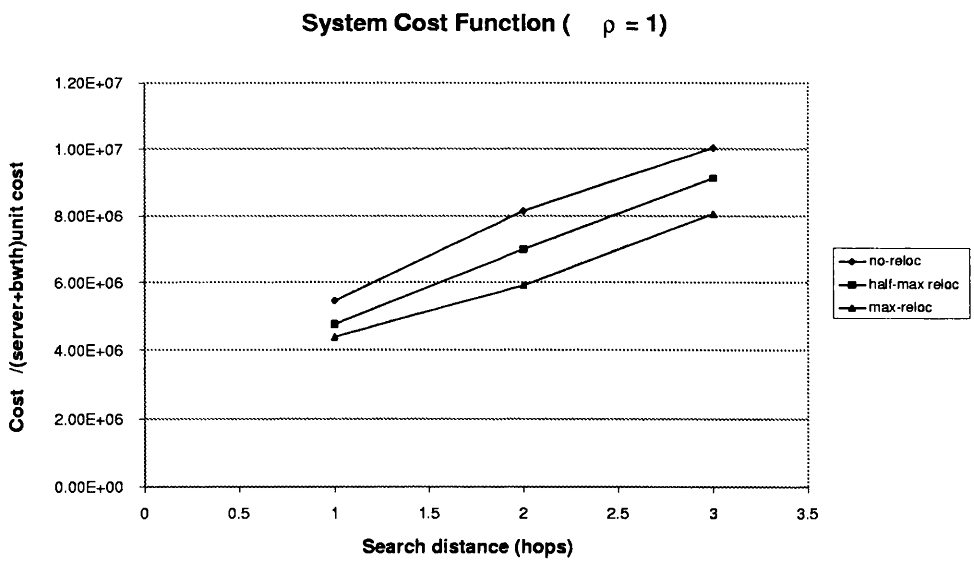

Figure 8.5: System cost function using servers of capacity 1

network. As the search range increases the local cluster capacity increases resulting in improved QoS.

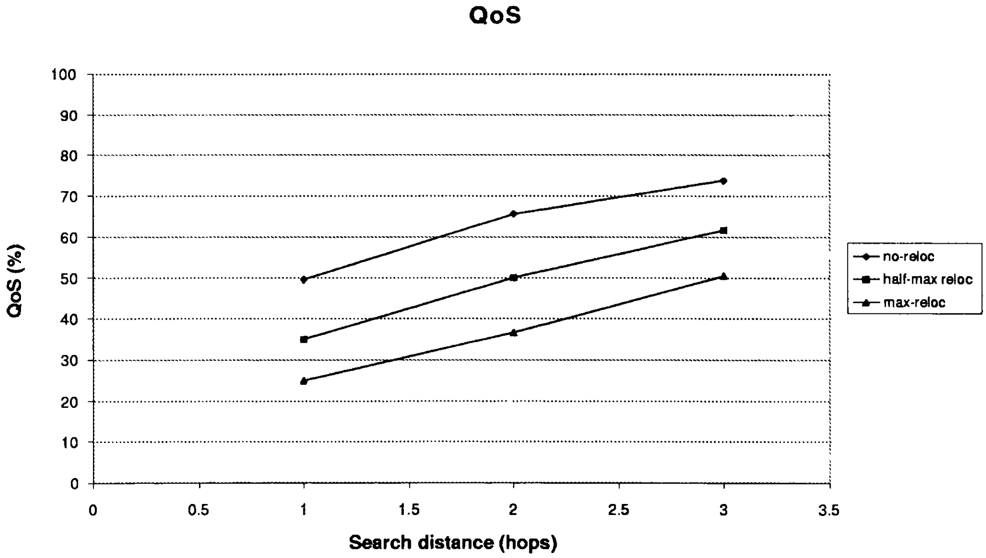

Figure 8.6: QoS achieved using servers of capacity 1

Using 3-movie server capacity Here we examine the effects of increasing the local search range from 1 to 3 hops, on a network comprising of servers with a capacity of 3 movies, for the non-relocating, half-relocating and full-relocating scenario.

In the 3-movie server capacity case, bandwidth usage decreases as search range increases from 1 to 2 hops for the full-relocating system, see *figure 8.7*. The decrease is of the order of 20%. In the half-relocating system a small decrease of bandwidth usage of the order of 5% is observed. In the non-relocating case, bandwidth usage increases steeply of the order of 150%.

As the search range increases even further from 2 hops to 3, bandwidth usage is kept at the same levels in the full-relocating system and is slightly less than in the half-relocating system, see *figure 8.7*. In the half-relocating system, bandwidth usage increases at a rate of 20%. In the non-relocating system, where bandwidth usage is the lowest, an increase of 40% is observed.

The connection success rate, increases as the search range increases from 1 to 2 hops, see *figure 8.8*. In the full-relocating system, where connection success rate is the highest, there is a steady increase of 10%. In the half-relocating

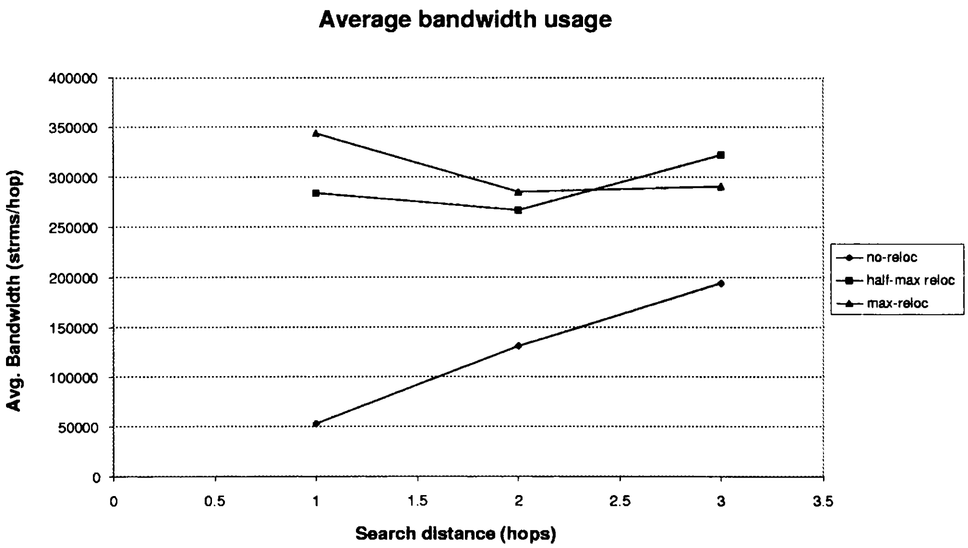

Figure 8.7: Average bandwidth usage using severs of capacity 3

system, with lower success rate, the rate is kept at the same levels. In the non-relocating system, where the success rate is the lowest, the increase rate is of the order of 10%.

Furthermore, as the search range increases from 2 to 3 three hops, the success connection rate continues to increase. In the full-relocating system, the success rate is the highest and the increase is of the order of 8%. In the half-relocating system the increase is of the order of 15% and in the non-relocating system with the lowest connection success rate the increase is 8% approximately, see *figure 8.8.*

The cut user percentage is almost zero for the half and full-relocating systems. In the non-relocating system the cut user percentage is small reaching 1.2×10^{-4} for the 2-hop search distance and dropping to a very small percentage of 4×10^{-5} for the 3-hop search distance, see *figure 8.9.*

The cost of the system is kept constant throughout the increase in search range in the full-relocating network and is the highest, see *figure 8.10.* In the half-relocating case, cost is kept constant in the 1 to 2 hops cases and increases by 10% in the 2 to 3 hops cases. In the non-relocating system, with the lowest

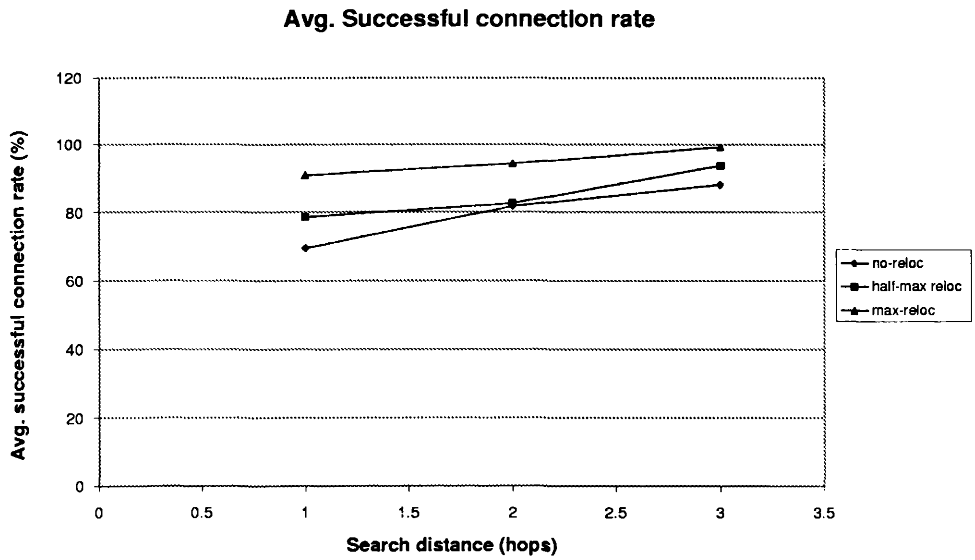

Figure 8.8: Average success connection rate using servers of capacity 3

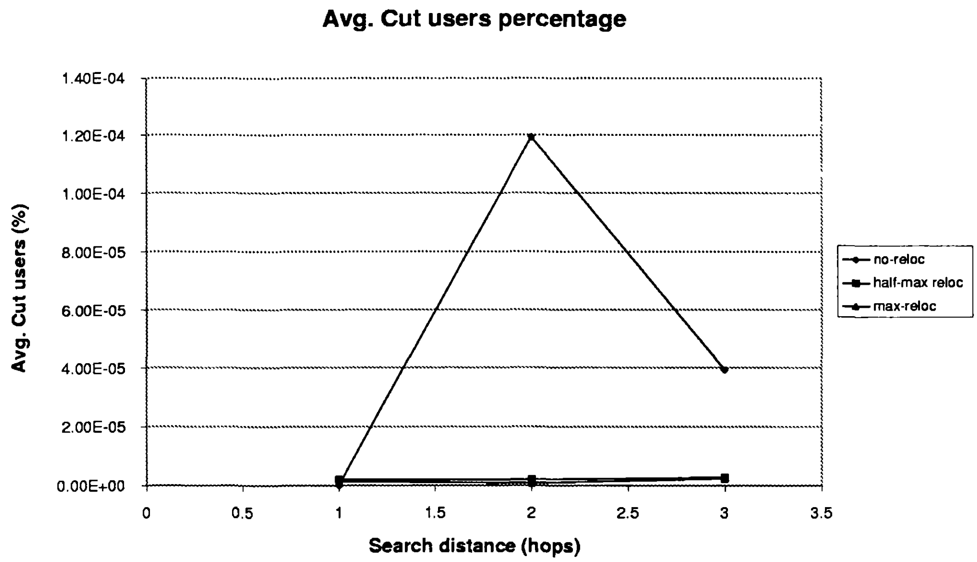

Figure 8.9: Average cut user percentage using servers of capacity 3

cost, cost increases by 10% in the 1 to 2 hops cases and further increases by 5% in the 2 to 3 hops cases.

The QoS offered by the system follows the same trends as the successful connection rate, see *figure 8.11.*

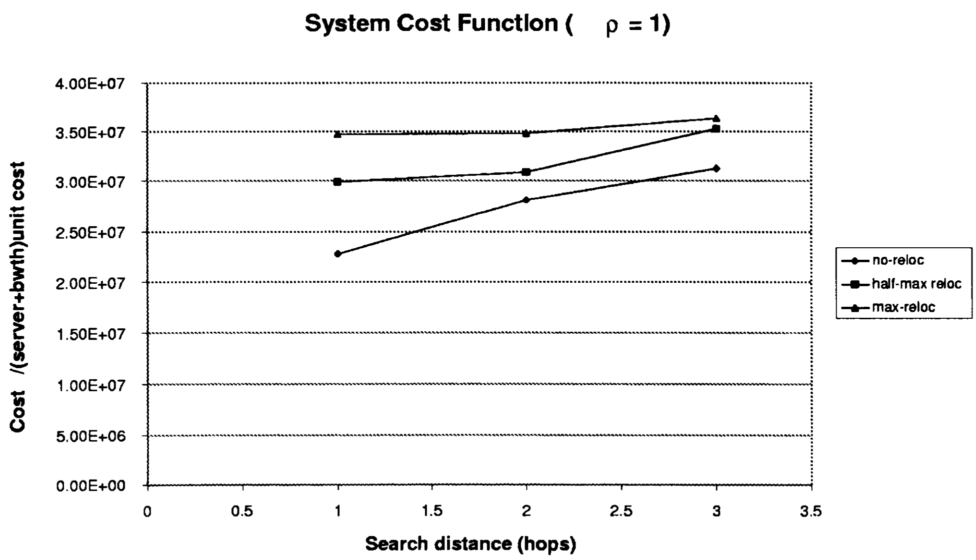

Figure 8.10: System cost function using servers of capacity 3

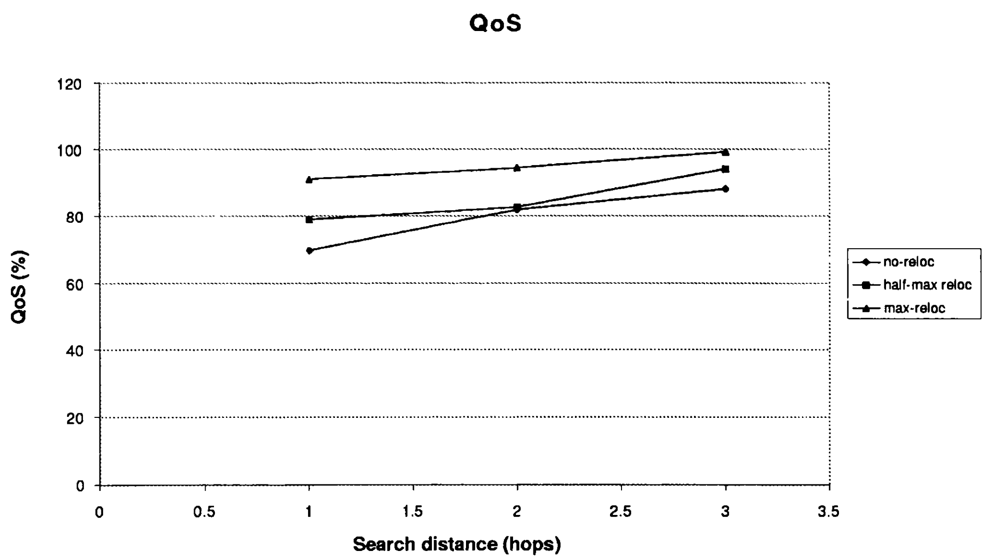

Figure 8.11: QoS achieved using servers of capacity 3

As the search distance increases from 1 to 2 and 3 hops the local server cluster capacity increases, as server capacity is 3 movies. This is reflected in the increase in QoS. Bandwidth usage initially decreases and then increases slightly but remains lower than in the 1-hop search range case.

Using 5-movie server capacity Here we examine the effects of increasing the local search range from 1 to 3 hops, on a network using servers of capacity 5 movies, for the non-relocating, half-relocating and full-relocating scenario.

In the 5-movie server capacity system, bandwidth usage decreases as search range increases from 1 to 2 hops, in the full-relocating system, see *figure 8.12.* The decrease is of the order of 10%. In the half-relocating system a small decrease of bandwidth usage of the order of 2% is observed. In the non-relocating case, bandwidth usage increases steeply in the order of 150%.

As the search range increases even further from 2 hops to three, bandwidth usage is kept at the same levels in the full-relocating system and is slightly less than in the half-relocating system. In the half-relocating system, bandwidth usage increases at a small rate of 5%. In the non-relocating system, where bandwidth usage is the lowest bandwidth usage increases at a rate of 40%.

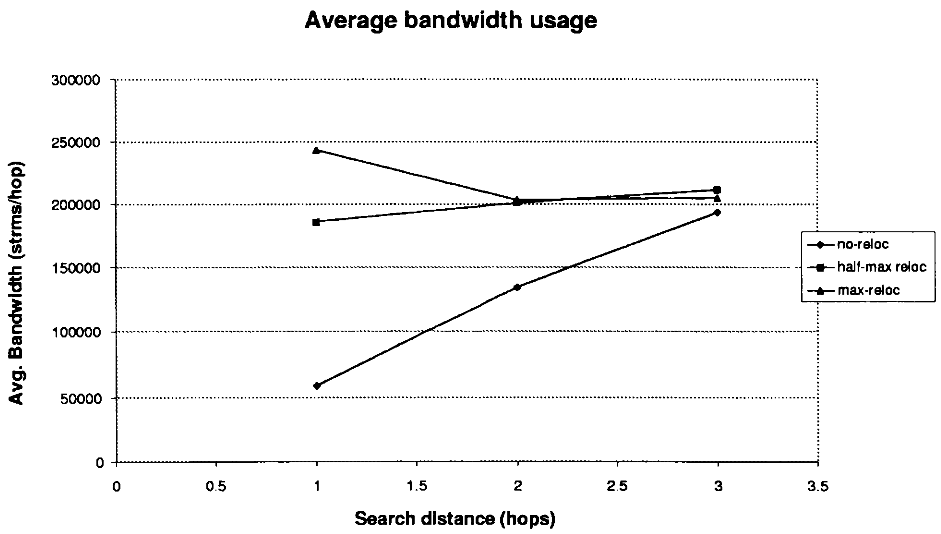

Figure 8.12: Average bandwidth usage using severs of capacity 5

The connection success rate, increases as the search range increases from 1 to 2 hops, see *figure 8.13.* In the full-relocating system, where connection success rate is the highest, there is a steady slight increase of 2%. In the half-relocating system, with lower success rate, there is a small increase of 5%. In

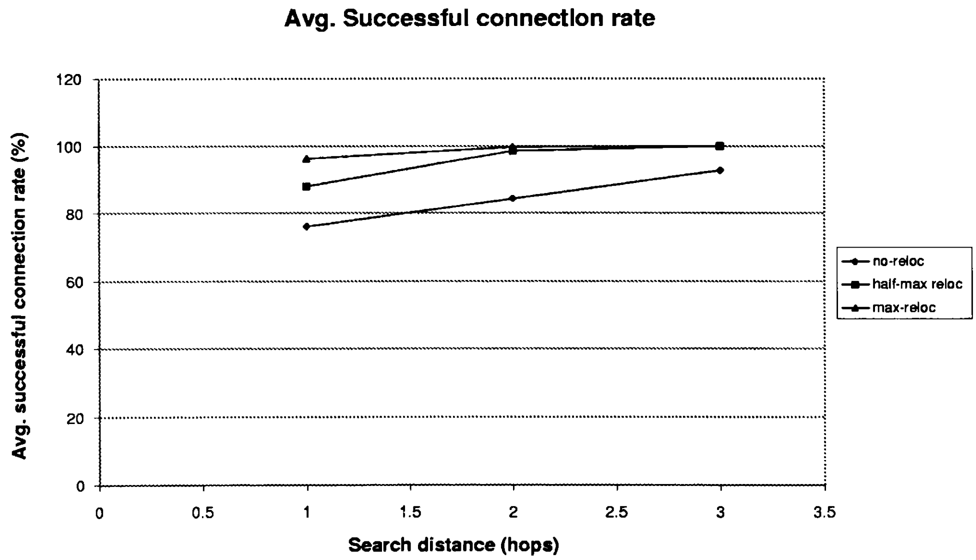

Figure 8.13: Average success connection rate using servers of capacity 5

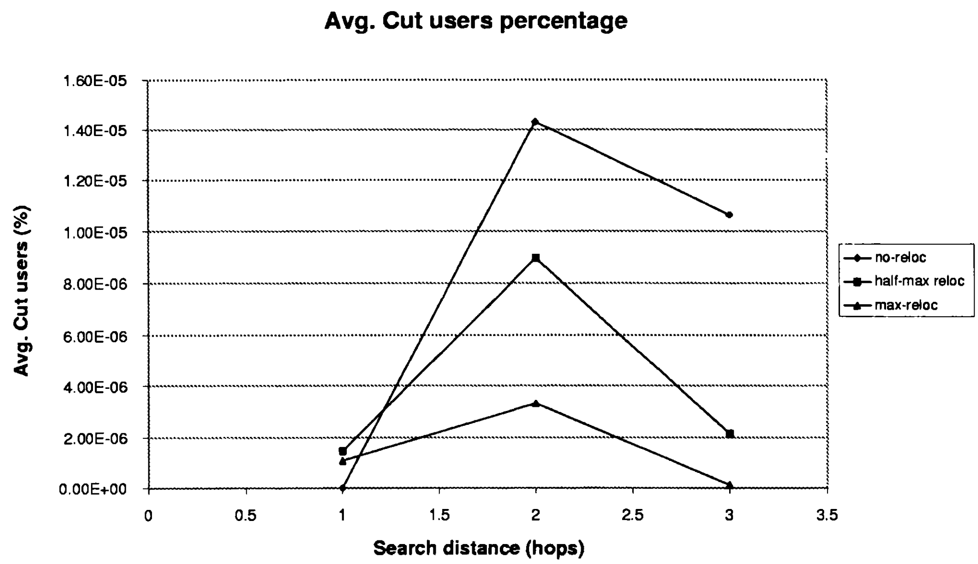

Figure 8.14: Average cut user percentage using servers of capacity 5

the non-relocating system, where the success rate is the lowest, the increase rate is of the order of 5%.

Furthermore, as the search range increases from 2 to 3 three hops, the success connection rate reaches 100% for both the relocating systems, with the full-relocating system performing slightly better. In the full-relocating system, the rate is the highest and kept constant. In the half-relocating system successful connection rate is almost constant, but slightly increasing and in the non-relocating system with the lowest connection success rate there is a 10% increase.

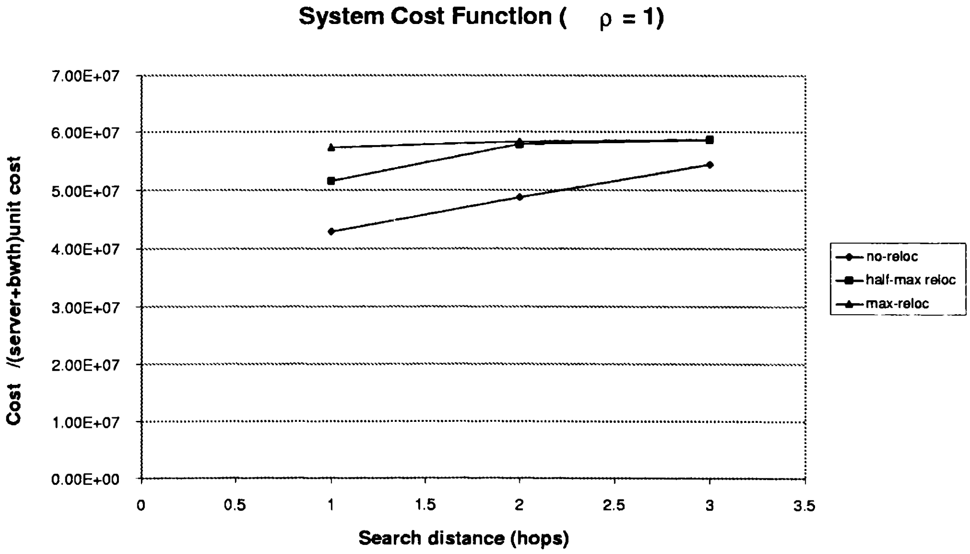

Figure 8.15: System cost function using servers of capacity 5

The cut user percentage is small for all systems, see *figure 8.14*. The highest cut user percentage is reached by the non-relocating system in the 2-hop case and is 1.4×10^{-5}, the half-relocating system follows with 9×10^{-6} and the full-relocating system with 3×10^{-6} all at the 2-hop case. The cut user percentage drops when using the 3-hop search distance at 1.1×10^{-5} for the non-relocating, 2×10^{-6} for the half-relocating and 0.01×10^{-6} for the full-relocating network.

The cost of the system is kept almost constant throughout the increase in search range in the full-relocating case and is the highest, see *figure 8.15.* In the half-relocating case, cost increases by 10% in the 1 to 2 hops cases and is kept constant in the 2 to 3 hops cases. In the non-relocating system, with the lowest cost, cost increases by 15% in the 1 to 2 hops cases and further increases by 8% in the 2 to 3 hops cases. The QoS offered by the system follows the same

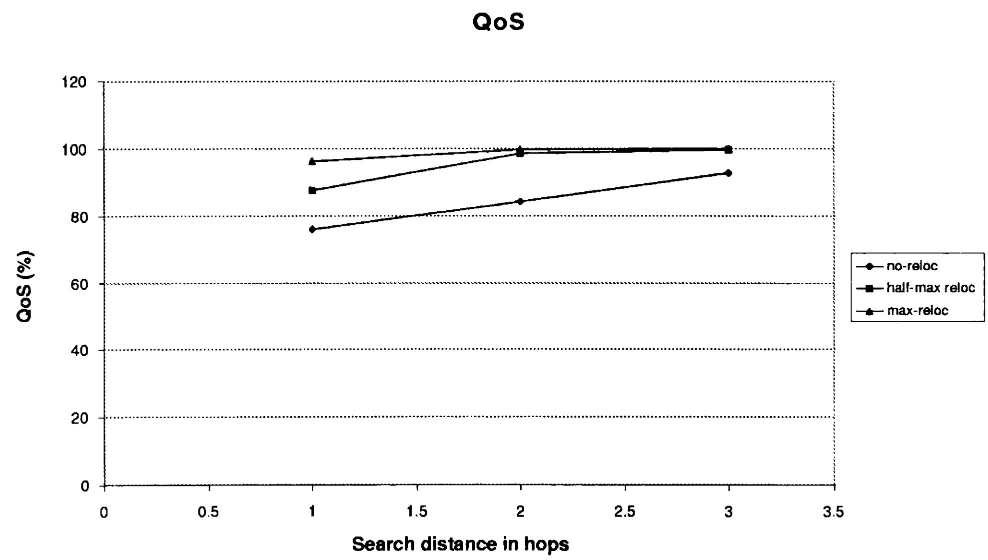

Figure 8.16: QoS achieved using servers of capacity 5

trends as the successful connection rate, see *figure 8.16.*

Here again the usefulness of the increased search range is demonstrated. Using servers of capacity 5 movies the capacity of the local cluster created increases thus prodiving better performance and QoS.

Using 8-movie server capacity Here we examine the effects of increasing the local search range from 1 to 3 hops, on an 8-movie server capacity network, for the non-relocating, half-relocating and full-relocating scenario.

In the 8-movie server capacity case, bandwidth usage decreases as search range increases from 1 to 2 hops, see *figure 8.17.* In the full-relocating system the decrease is of the order of 15%. In the half-relocating system bandwidth usage is almost kept constant. In the non-relocating case, bandwidth usage increases steeply at a rate of 86%.

As the search range increases even further from 2 hops to three, bandwidth usage increases in all three systems. In the full and half-relocating systems bandwidth usage increases at the same rate of 20%. In the non-relocating system, bandwidth usage is the lowest at the 2-hop case and becomes the highest in the 3-hop case, increase.

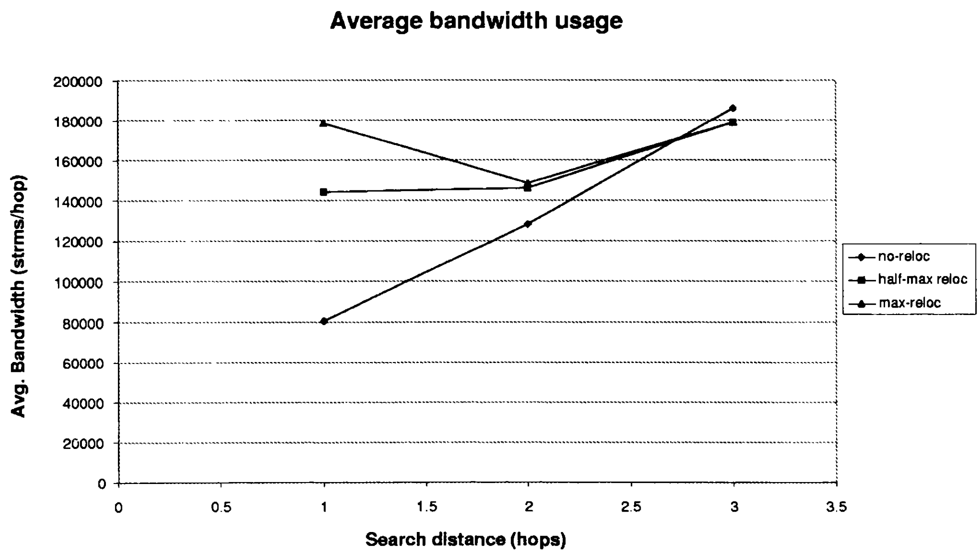

Figure 8.17: Average bandwidth usage using severs of capacity 8

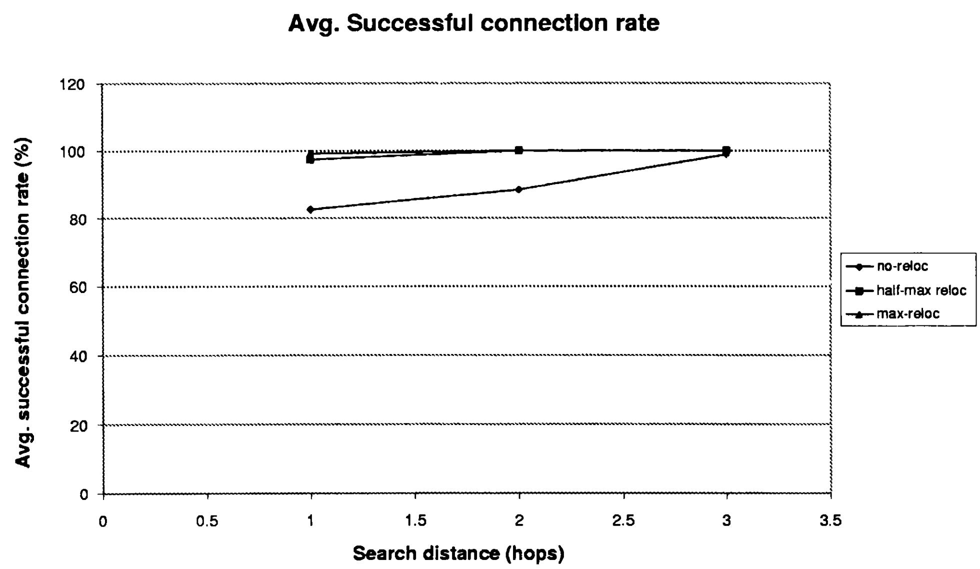

Figure 8.18: Average success connection rate using servers of capacity 8

The connection success rate, increases as the search range increases from 1 to 2 hops, see *figure 8.18*. In the half and full-relocating system, where connection

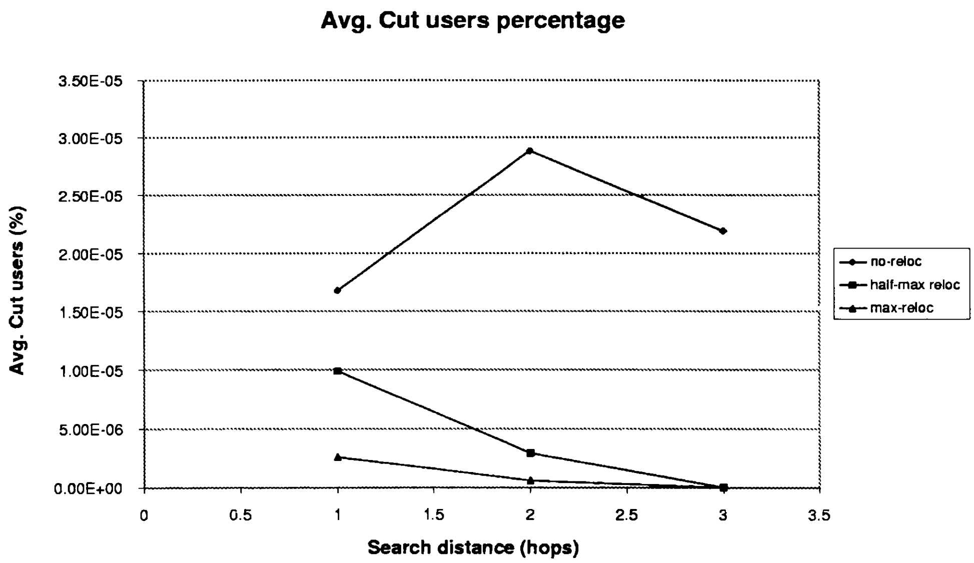

Figure 8.19: Average cut user percentage using servers of capacity 8

success rate is the highest, there is a steady very slight increase, with the full-relocating system performing slightly better. In the non-relocating system, with lowest success rate, there is a small increase of 8%.

Furthermore, as the search range increases from 2 to 3 three hops, the success connection rate reaches 100% for both the relocating systems, with the full-relocating system performing slightly better. In the full-relocating system, the rate is the highest and kept constant. In the half-relocating system successful connection rate is almost constant, but slightly increasing and in the non-relocating system with the lowest connection success rate the increase is of the order of 10%.

The cut user percentage is small for all systems, see *figure 8.19.* The highest cut user rate is reached by the non-relocating system in the 1-hop case is 1.75×10^{-5}, in the 2-hop case it increases to 2.75×10^{-5} and in the 3-hop case drops to 2.25×10^{-5}. The half-relocating system follows with 1×10^{-5} at the 1-hop case and the full-relocating system with 2.5×10^{-6} the 2-hop case.

The cost of the system is kept almost constant throughout the increase in search range in the half and full-relocating cases, being higher in the later, see

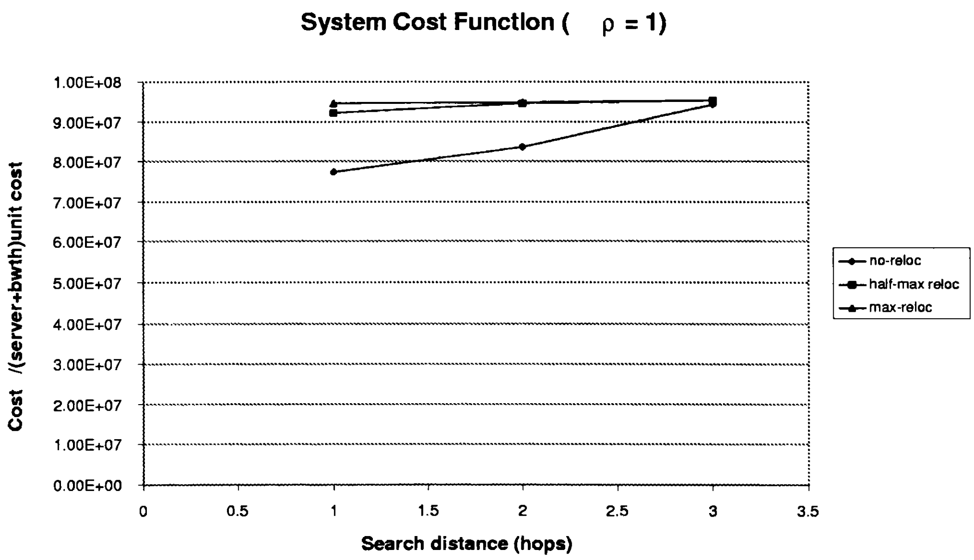

Figure 8.20: System cost function using servers of capacity 8

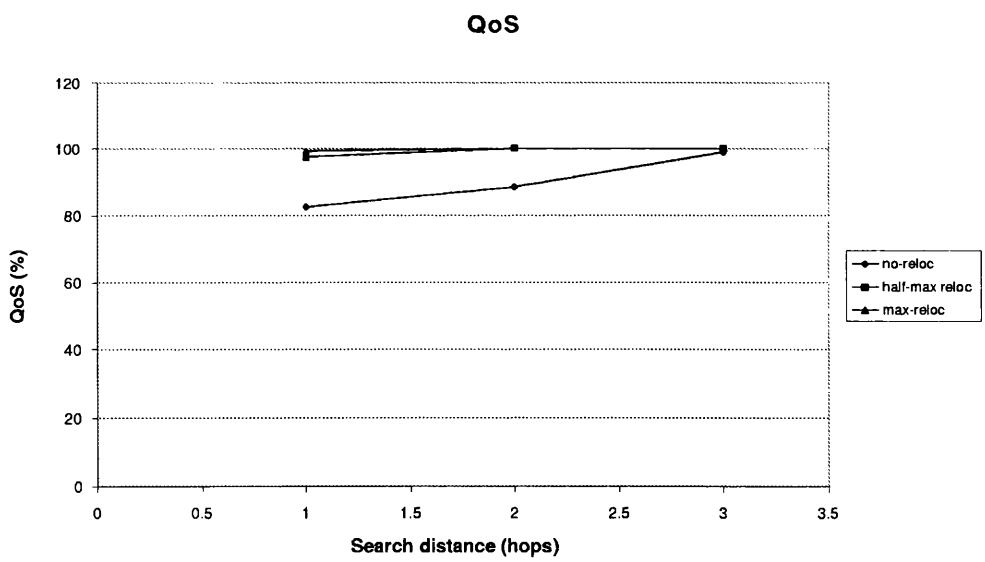

Figure 8.21: QoS achieved using servers of capacity 8

figure 8.20. In the non-relocating system, with the lowest cost, cost increases by 5% in the 1 to 2 hops cases and further increases by 8% in the 2 to 3 hops cases.

The QoS offered by the system follows the same trends as the successful connection rate, see *figure 8.21.*

Using servers of capacity 8 movies and a search distance of 2 hops is enough to provide almost perfect QoS, increasing the search range a slightly better performance is provided at the cost of increasing bandwidth usage. Overall system performance is very good.

Using 10-movie server capacity Here we examine the effects of increasing the local search range from 1 to 3 hops, on a 10-movie server capacity network, for the non-relocating, half-relocating and full-relocating scenario.

In the 10-movie server capacity case, bandwidth usage is almost kept constant as search range increases from 1 to 2 hops, in the full-relocating system. In the half-relocating system bandwidth usage increases slightly by 4%. In the non-relocating case, bandwidth usage increases steeply at a rate of 117%, see *figure 8.22.*

As the search range increases even further from 2 hops to three, bandwidth usage increases in all three systems, at the same rate of 35%.

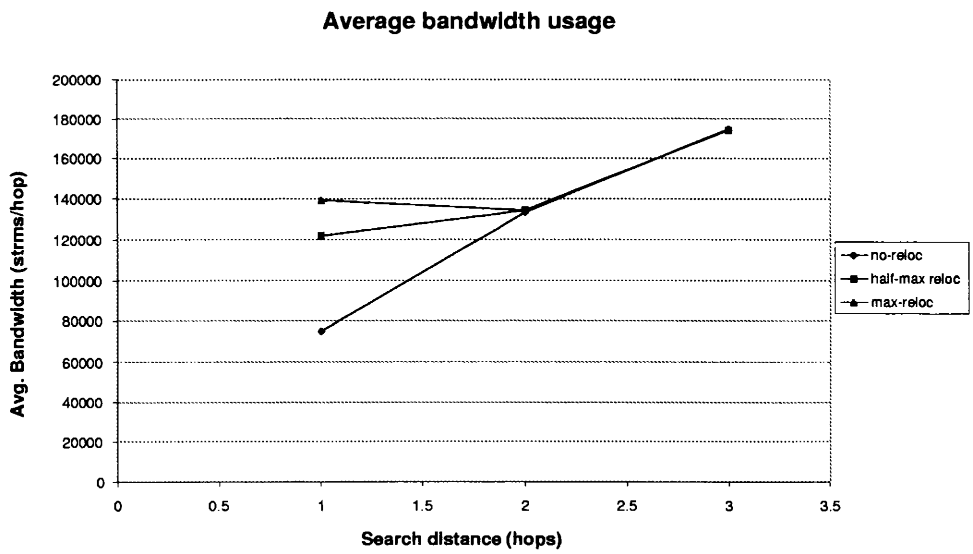

Figure 8.22: Average bandwidth usage using severs of capacity 10

The connection success rate, increases as the search range increases from 1 to

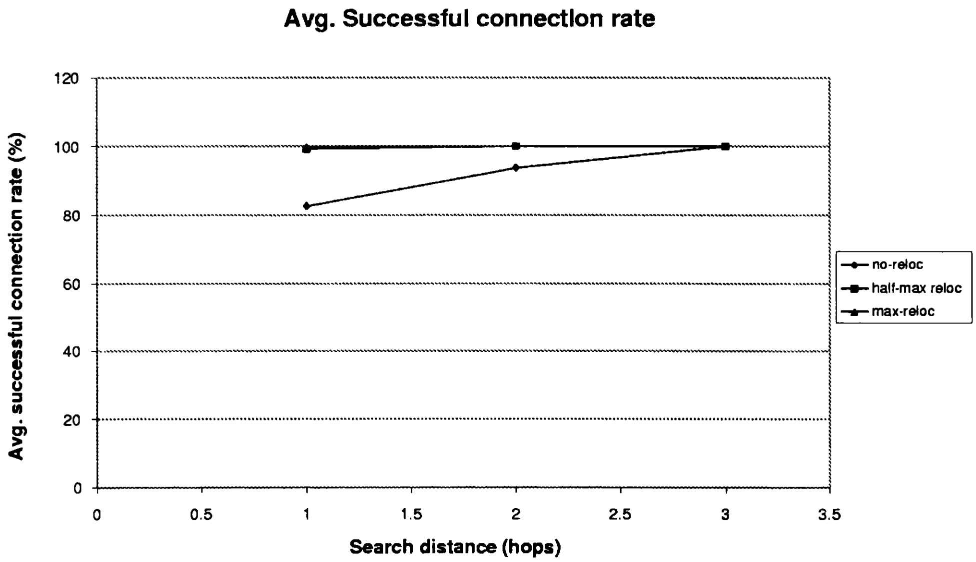

Figure 8.23: Average success connection rate using servers of capacity 10

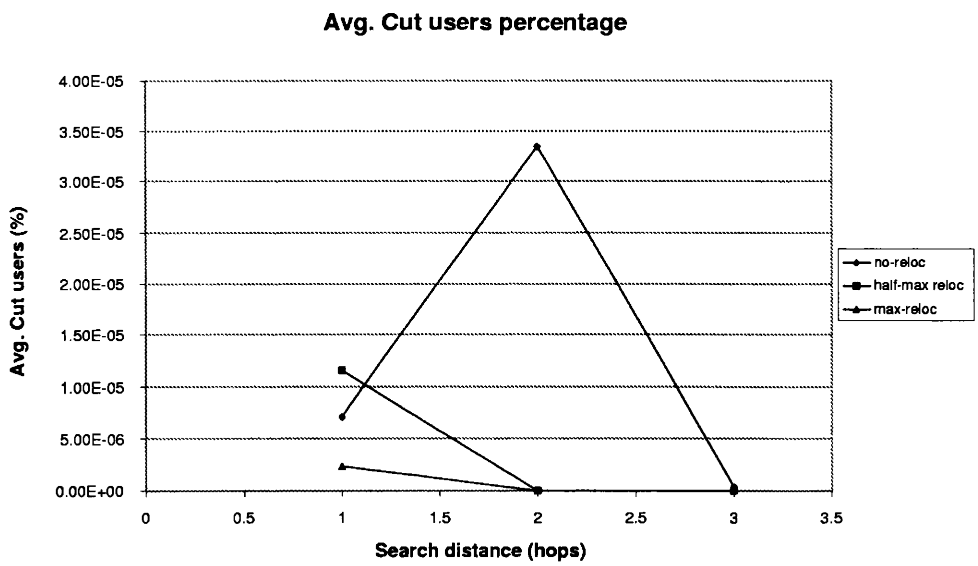

Figure 8.24: Average cut user percentage using servers of capacity 10

2 hops, see *figure 8.23*. In the half and full-relocating system, where connection success rate is the highest, connection success rate is almost constant at 100%, with the full-relocating system performing slightly better. In the non-relocating system, with lowest success rate, there is a small increase of 12%.

Furthermore, as the search range increases from 2 to 3 three hops, the success connection rate remains at 100% for both the relocating systems. In the non-relocating system with the lowest connection success rate increases at the rate of 10%.

The cut user percentage is small for all systems, see *figure 8.24*. The highest cut user rate is reached by the non-relocating system in the 2-hop case and is 3.4×10^{-5}, the half-relocating system follows with 1.75×10^{-5} at the 1-hop case and the full-relocating system with 2.5×10^{-6} at the 2-hop case.

The cost of the system is kept almost constant throughout the increase in search range in the half and full-relocating cases, being higher in the latter, see *figure 8.25*. In the non-relocating system, with the lowest cost, cost increases by 10% in the 1 to 2 hops cases and further increases by 8% in the 2 to 3 hops cases.

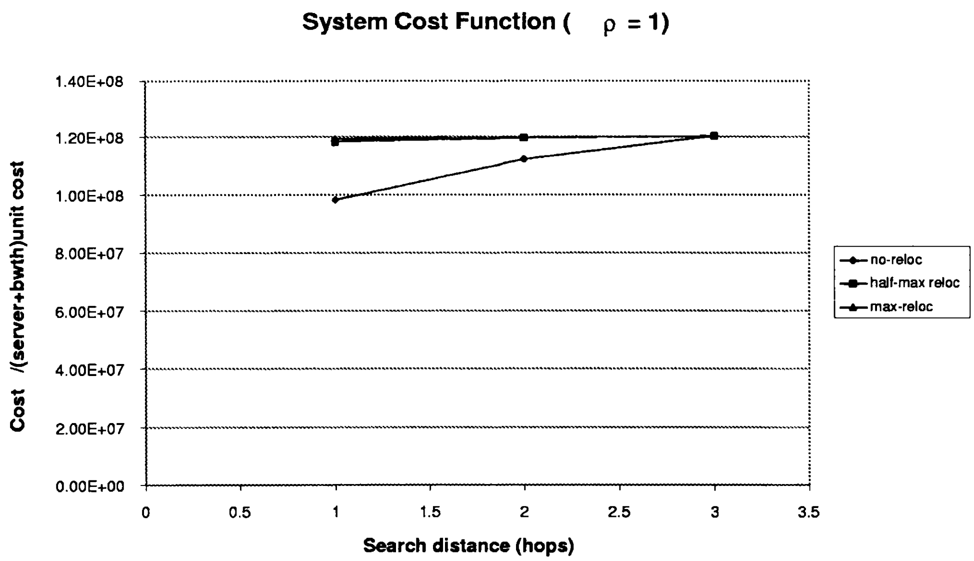

Figure 8.25: System cost function using servers of capacity 10

The QoS offered by the system follows the same trends as the successful connection rate, see *figure 8.26.*

The server capacity of 10 movies is enough for the system to provide service to all users requesting it.

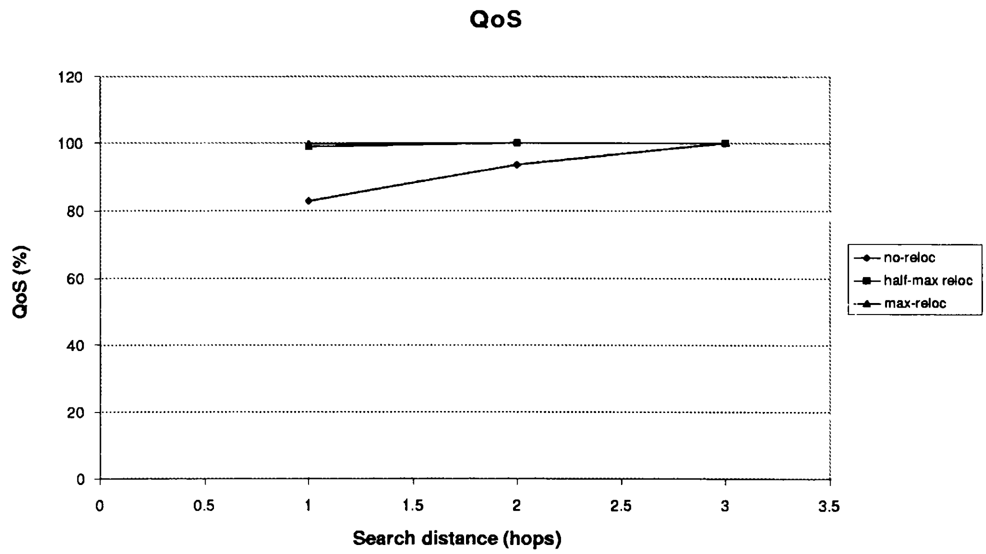

Figure 8.26: QoS achieved using servers of capacity 10

As server capacity increases to 10 movies the relocating systems perform almost perfectly using the same amount of bandwidth as the non-relocating one. The increase in cost is due to the extra connected users. The QoS value is very good too due to the very small number of cut users.

Using 15-movie server capacity Here we examine the effects of increasing the local search range from 1 to 3 hops, on a 15-movie server capacity network, for the non-relocating, half-relocating and full-relocating scenario.

In the 15-movie server capacity case, bandwidth usage increases as search range increases from 1 to 2 hops, in both half and full-relocating systems, see *figure 8.27*. In both the full and half-relocating systems bandwidth usage increases at the rate of 20%. In the non-relocating case, bandwidth usage increases more steeply at a rate of 45%.

As the search range increases even further from 2 hops to three, bandwidth usage increases in all three systems, at the same rate of 20%.

The connection success rate, is very high in both cases in the 1 and 2 hop cases, see *figure 8.28*. In the half and full-relocating system, where connection

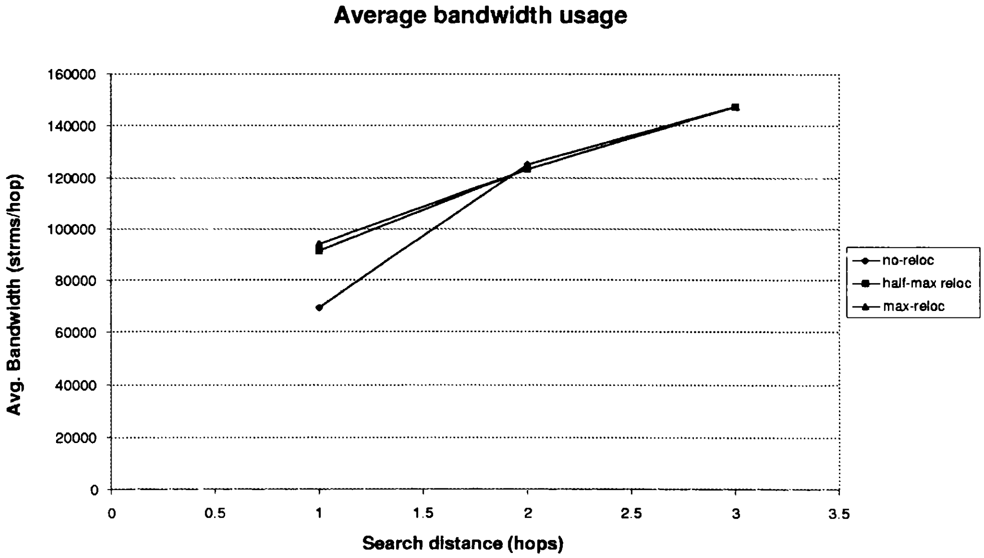

Figure 8.27: Average bandwidth usage using severs of capacity 15

success rate is the highest, connection success rate is almost constant at 100%, with the full-relocating system performing slightly better. In the non-relocating system, with lowest success rate, there is an increase of 14%.

Furthermore, as the search range increases from 2 to 3 three hops, the success connection rate remains at 100% for all three systems.

The cut user percentage is small for all systems, see *figure 8.29.* The highest cut user rate is reached by the non-relocating system in the 1-hop case and is 2.75 10^{-5}, the half-relocating system follows with 5×10^{-6} at the 1-hop case and the full-relocating system with 2×10^{-6} at the 1-hop case.

The cost of the system is kept almost constant throughout the increase in search range in the half and full-relocating cases, being higher in the later, see *figure 8.30.* In the non-relocating system, with the lowest cost, cost increases by 15% in the 1 to 2 hops cases and is kept constant in the 2 to 3 hops cases.

Here as server capacity increases further, the system has enough resources to accommodate all users within the local server cluster. This is reflected in the high QoS for all systems. Cut users are dropping to zero as the local server cluster radius increases to 3 hops.

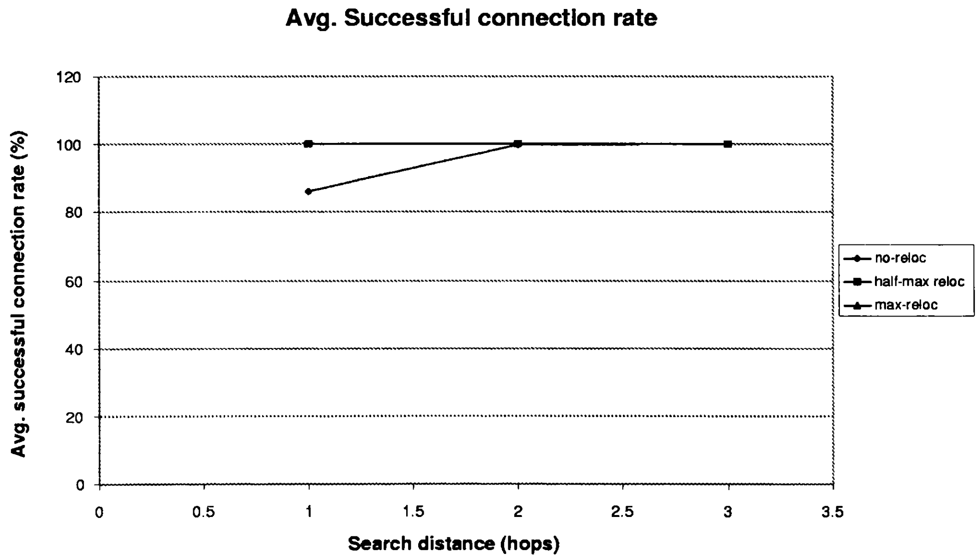

Figure 8.28: Average success connection rate using servers of capacity 15

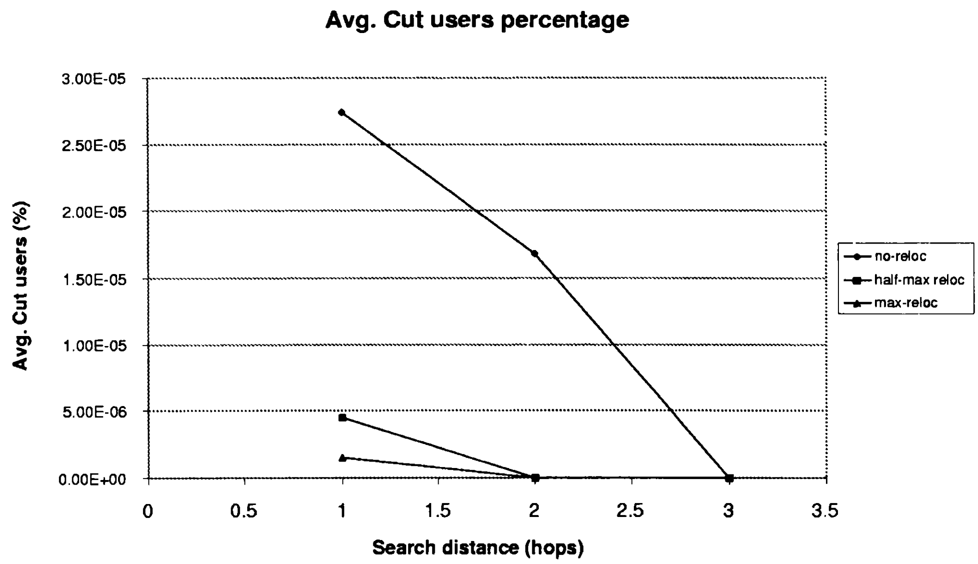

Figure 8.29: Average cut user percentage using servers of capacity 15

The QoS offered by the system follows the same trends as the successful connection rate, see *figure 8.31*.

The relocating systems perform better when server capacity exceeds 3 movies. Then relocation is most effective when enough resources are available for the

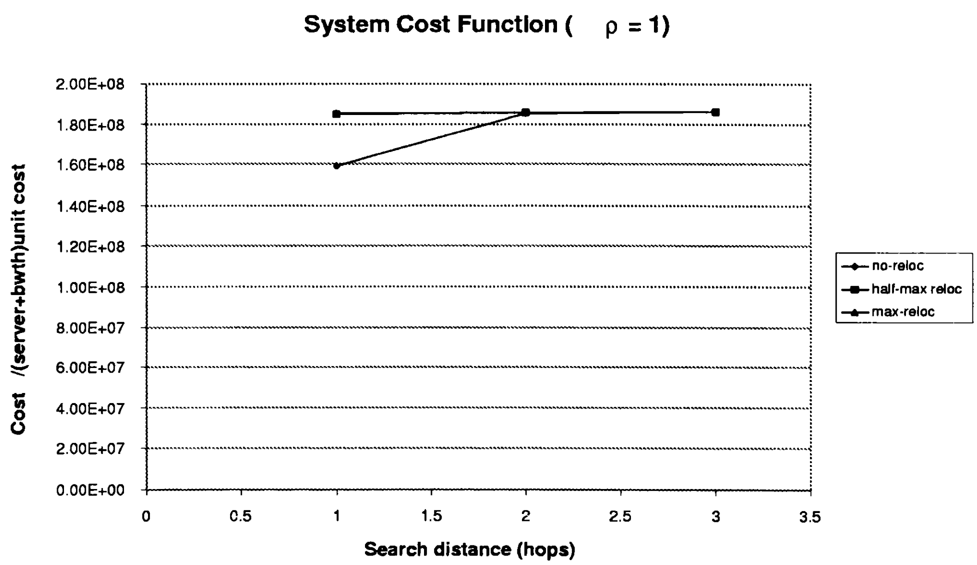

Figure 8.30: System cost function using servers of capacity 15

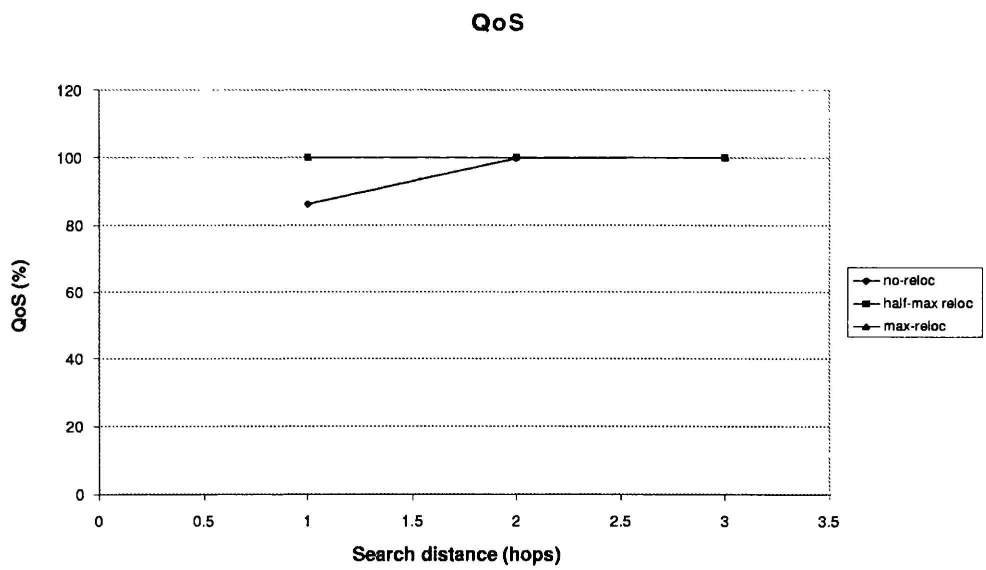

Figure 8.31: QoS achieved using servers of capacity 15

operation to function properly. Although bandwidth usage is comparable in the high server capacity cases, service performance is further improved when using relocation. Cost increases linearly in the low capacity cases of 1,3 and 5 movies and at similar constant levels in the higher server capacity cases. Relocation is a method, which can be used to improve the usage of system resources and improve QoS. The service provider should keep in mind that the cost function given here is only a generalised one and can change depending on real component prices.

System performance with increasing server capacity and constant search distance

Next we examine the effects of increasing the server capacity from 1 to 15 movies, for the non-relocating, half-relocating and full-relocating scenario keeping constant the local cluster search distance.

Using 1-hop search distance Here we examine the effects of increasing the local server capacity and using a local search range of 1-hop, for the non-relocating, half-relocating and full-relocating scenario.

As local server capacity is one movie, the local server is capable of accommodating only a small percentage of the total demand for the 100-movie library. The control mechanism gives priority to the very popular movies and tries to set up as many of these as possible. The rest of the movies are started on remote servers and some are relocated later on. Bandwidth usage is moderate as is user connection rate, see *figure 8.32* and there are no cut users *figure 8.34*, as only the more popular movies are started and the least popular ones are used for a substantial number of users.

As local server capacity becomes three movies, the local server is capable of accommodating only a small percentage of the total demand for the 100-movie library. The control mechanism is relocating heavily in order to improve utilisation of system resources. This produces the huge increase of bandwidth usage and consequently the full-relocation case demonstrates more bandwidth usage, see *figure 8.32*. On the other hand the system does perform better on the account of servicing more users as shown in the connection success graph in *figure 8.33*, but the increase in connection rate does not follow the steep increase in bandwidth usage because each server tries to accommodate as many users

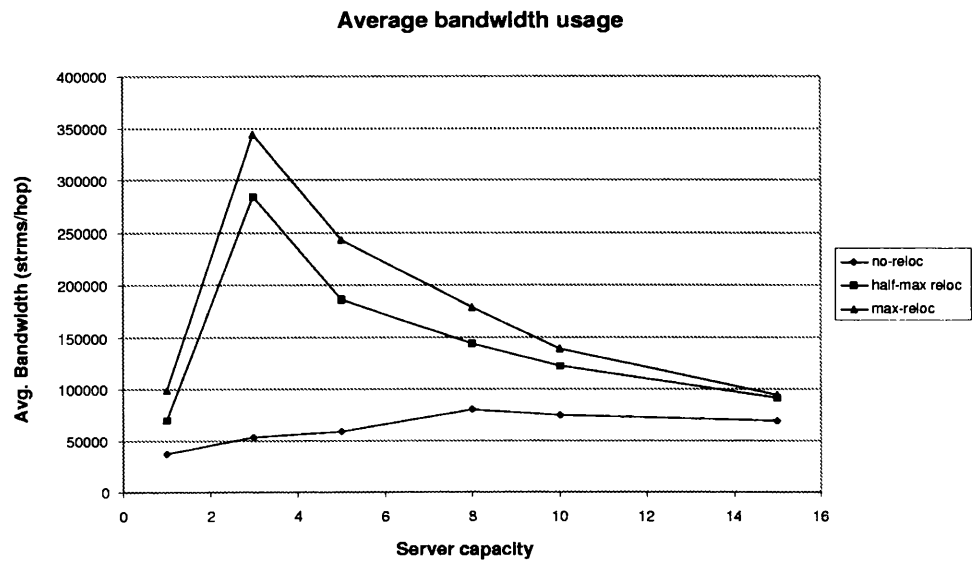

Figure 8.32: Average bandwidth usage using search distance of 1 hop

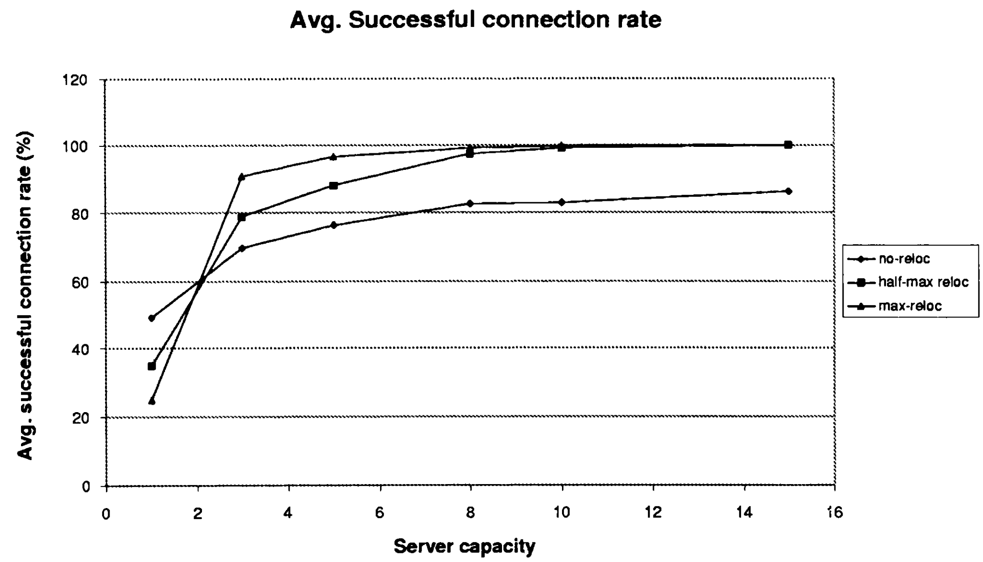

Figure 8.33: Average success connection rate using search distance of 1 hop

as possible locally. In the cut user rate, see *figure 8.34*, all systems perform similarly since there is a user threshold before a movie is cut.

As local server capacity increases to 5 movies system performance improves on all counts. Bandwidth usage drops significantly, see *figure 8.32*. Now the

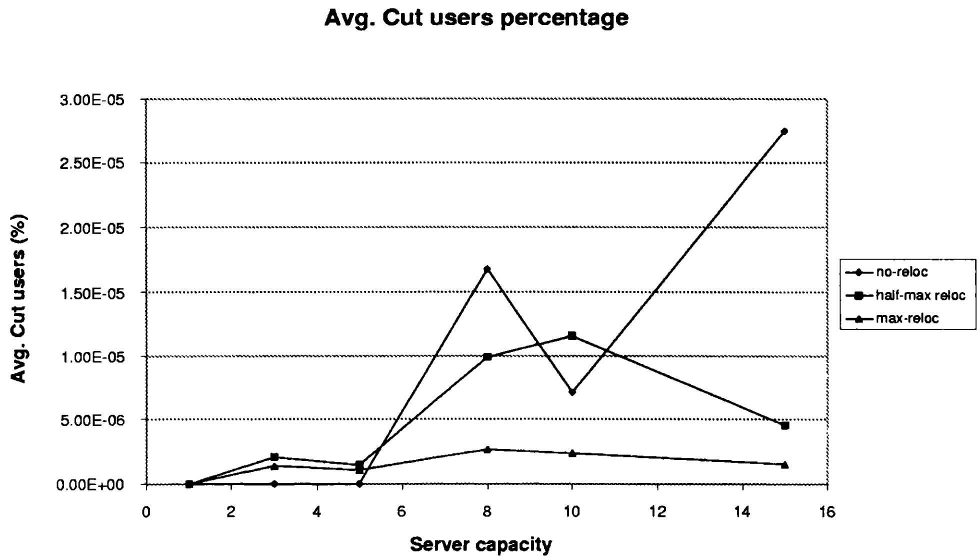

Figure 8.34: Average cut user percentage using search distance of 1 hop

local server is able to hold 5 of the most popular movies and the local cluster of servers within 1 hop from it can hold another 20 movies totalling the 25 most popular movies of the library. User success rate increases, see *figure 8.32*. The full relocating scenario performs better, since a small number of users can make way for another more popular movie and due to the increased server size this is now possible is most attempts. In the half relocation case some relocation attempts fail as the required movies are beyond the scope of the search. The number of cut users is small, see *figure 8.34*, as the more popular movies are allocated most places.

In the 8-movie local server capacity, system performance is improved. Now the local cluster capacity increases to 40 movies including some of the less popular ones. Bandwidth usage drops, see *figure 8.32*, but not at the same rate as before since we are connecting less popular movies to the local cluster. Here the full relocating case performs better for the same reasons as in the previous case.

In the 10-movie local server capacity, the cluster capacity reaches half of the movies provided by the library including all movies but the least popular ones.

Now the relocating scenarios perform similarly as the required movie is within the search range of the half relocation case, see *figure 8.32*. There is a number of cut users since the flexibility of relocation is not available, see *figure 8.34*.

In the 15-movie local server capacity, the cluster capacity increases even further to include some of the least popular movies, this is demonstrated by the small increase in the user connection rate, see *figure 8.33*. Bandwidth usage reduction is due to the same reason, see *figure 8.32*. Now that quite a few of the least popular movies are played more fulfil the requirements of cutting down a movie to replace it by a more popular one, as shown in *figure 8.34*.

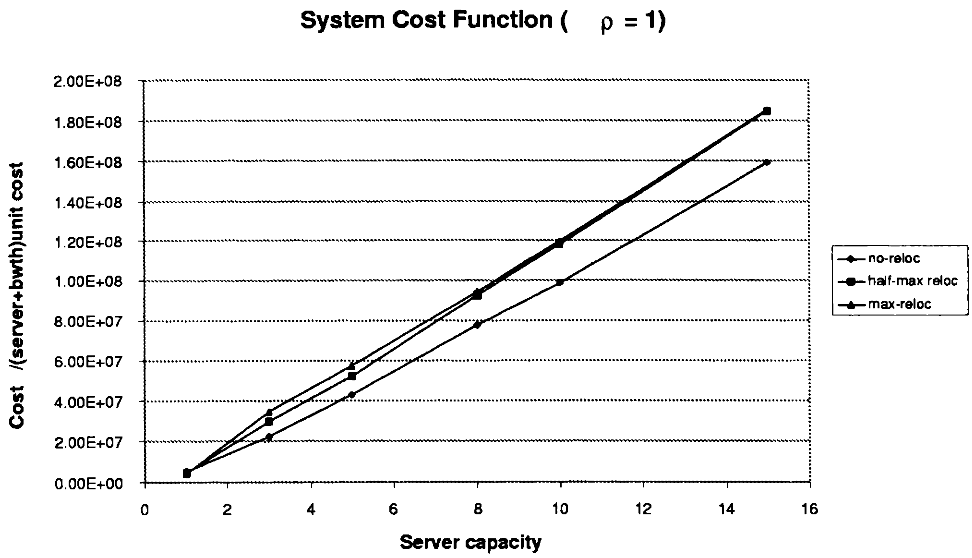

Figure 8.35: System cost function using search distance of 1 hop

In the case of using 1-hop search distance, bandwidth usage rises steeply as server capacity increases from 1 to 3 movies, in the cases of half and full range relocation, as shown in *figure 8.33*. The increase is of the order of 6 times more bandwidth. In the non-relocation case the bandwidth increase is of the order of 10%. As server capacity increases even further, bandwidth usage falls at a less steep rate. As capacity increases from 3 to 5 movies the bandwidth usage is reduced by a third, in the cases where relocation is used. The non-relocating network uses almost the same amount of bandwidth. As server capacity is

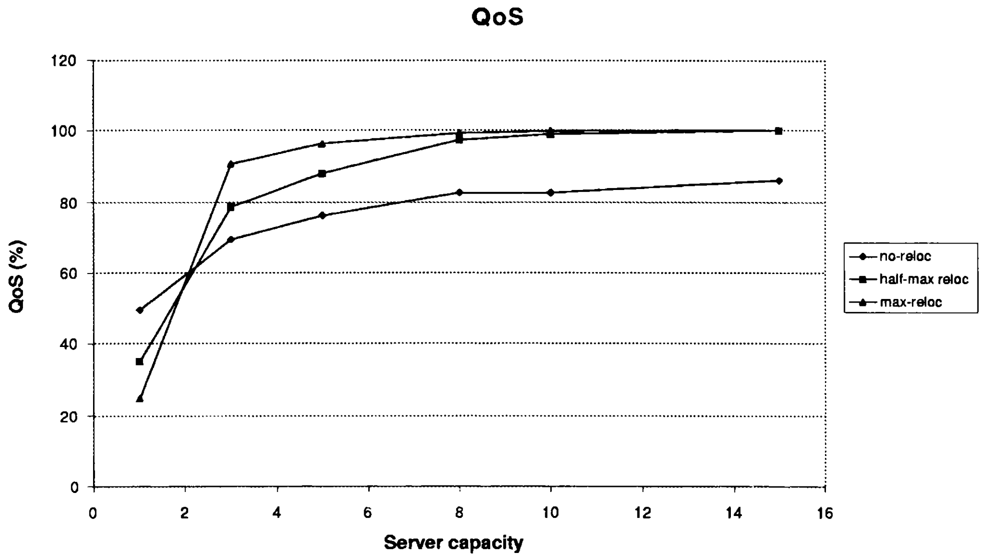

Figure 8.36: QoS achieved using search distance of 1 hop

increased from 5 to 10 movies, bandwidth usage is further decreased by one third, in the relocating systems. In the non-relocating system, there is an increase of 50%. Finally, as server capacity is increased to its maximum of 15 movies, bandwidth usage decreases further reaching almost the same levels of the one-movie server capacity cases, in both the relocating systems. In the non-relocating system bandwidth usage is kept at the same level.

The average connection success rate is almost tripled in both relocating cases, see *figure 8.32*, as server capacity is increased from one to three movies. In the non-relocating case there is an increase, but is not that steep, it is of the order of 40%. As server capacity is increased further from 3 to 8 movies, the rate of successful connections is increased by 10%-20%, in all networks. As server capacity is increased to its maximum of 15, successful connection rate is increased but only slightly in the rate of 3%, in all cases. In the relocating cases 100% successful connection rate is reached whereas in the non-relocating case the maximum reached is around 80%.

The percentage of cut users is quite small in all cases, see *figure 8.34*. The highest values are encountered in the non-relocating case of 2.75×10^{-5}. The lowest values are encountered in the full-relocating network. The system cost rises linearly as server capacity is increased in all cases, see *figure 8.35.* The QoS plot pattern follows the successful connection rate as cut user percentage is low, see *figure 8.36.*

Using 2-hop search distance Here we examine the effects of increasing the local server capacity and using a local search range of 2 hops, for the non-relocating, half-relocating and full-relocating scenario.

As local server capacity is one movie, the local server is capable of accommodating only a small percentage of the total demand for the 100-movie library and the server cluster capacity is 13 movies. The control mechanism gives priority to the very popular movies and tries to set up as many of these as possible. The rest of the movies are started on remote servers and some are relocated later on. Bandwidth usage is moderate but increased now that the search distance has increased to 2 hops, see *figure 8.37.* The user connection rate is increased as well since the more popular movies are organised on a wider server cluster, as shown in *figure 8.38.* There are no cut users, see *figure 8.39* as only the more popular movies are started and the least popular ones are used for a substantial number of users.

As local server capacity is three movies, the local server is capable of accommodating only a small percentage of the total demand for the 100-movie library, but the local server cluster is capable of providing most of the movies in the library. The control mechanism is relocating some of the movies in order to better utilise system resources. This produces the increase of bandwidth usage observed in *figure 8.37* and consequently the full-relocation case demonstrates a slight increase in bandwidth usage. This increase is significantly less, than in the 1-hop search case and so is the gap between the full and half relocation cases, due to the increased server size. On the other hand system performance reaches 80% in all cases as shown in the connection success graph in *figure 8.38*, the increase in connection rate does not follow the increase in bandwidth usage,

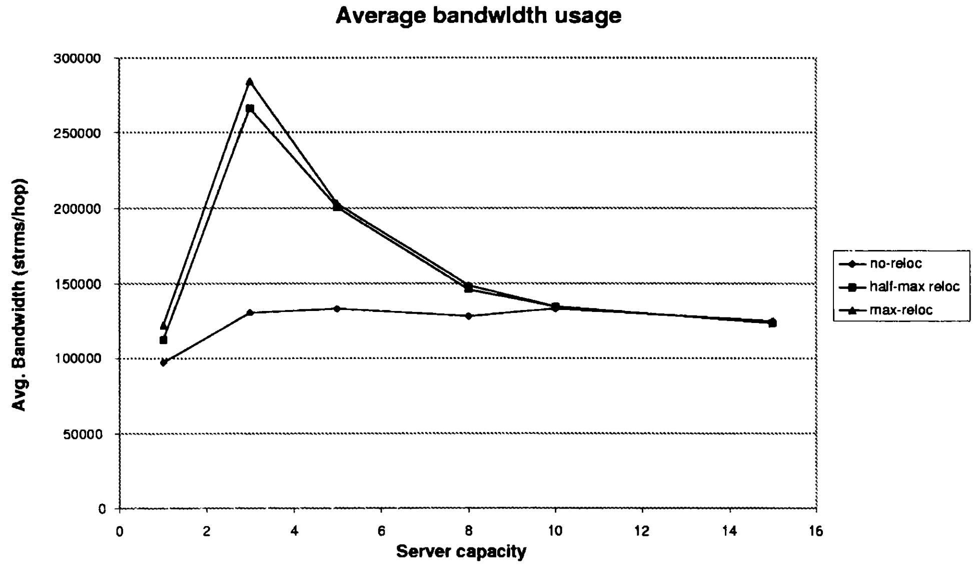

Figure 8.37: Average bandwidth usage using search distance of 2 hops

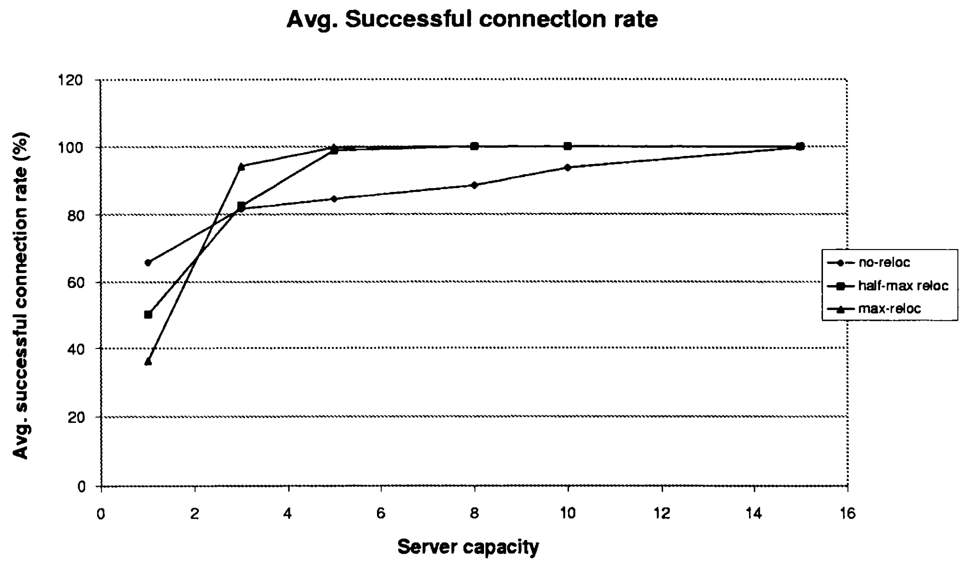

Figure 8.38: Average success connection rate using search distance of 2 hops

it is similar though, since each server tries to accommodate as many users as possible firstly locally and secondly on the local cluster. In the cut user rate the non-relocating case cuts the most users as it lacks the flexibility of relocating less popular movies, it is forced to discontinue them, see *figure 8.39*.

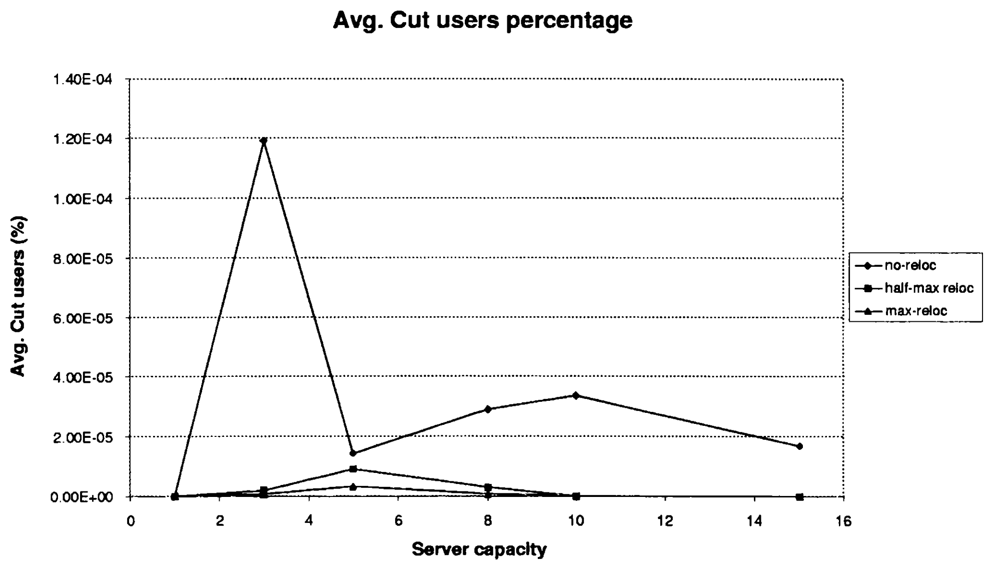

Figure 8.39: Average cut user percentage using search distance of 2 hops

As local server capacity increases to 5 movies system performance improves on all counts. Bandwidth usage drops significantly, see *figure 8.37*. Now the local server is able to hold 5 of the most popular movies and the local cluster of servers within 2 hops from it can hold another 60 movies. The total of 60 movies includes all the most popular ones and 15 of the least popular ones. User success rate increases, since well above 80% of the users can connect within the local cluster, see *figure 8.38*. The full relocating scenario performs better, since a small number of users can make way for another more popular movie and due to the increased server size this is now possible for most relocation attempts. In the half relocation case some relocation attempts fail as the required movies are beyond the scope of the search. Compared to the 1-hop search range case, the non-relocating scenario consumes more bandwidth but performs better, see *figure 8.37*, since due to the increased search range more bandwidth is used to access the more popular movies which are now within the user's reach. The half-relocating case consumes the same bandwidth and performs similarly as before, since there are still failed relocation attempts due to the restricted scope of the search range. The full-relocating case now uses less bandwidth and performs very well (95%) since increased cluster capacity and full-relocation allow the placement of the more popular movies on the best access points for the user pool.

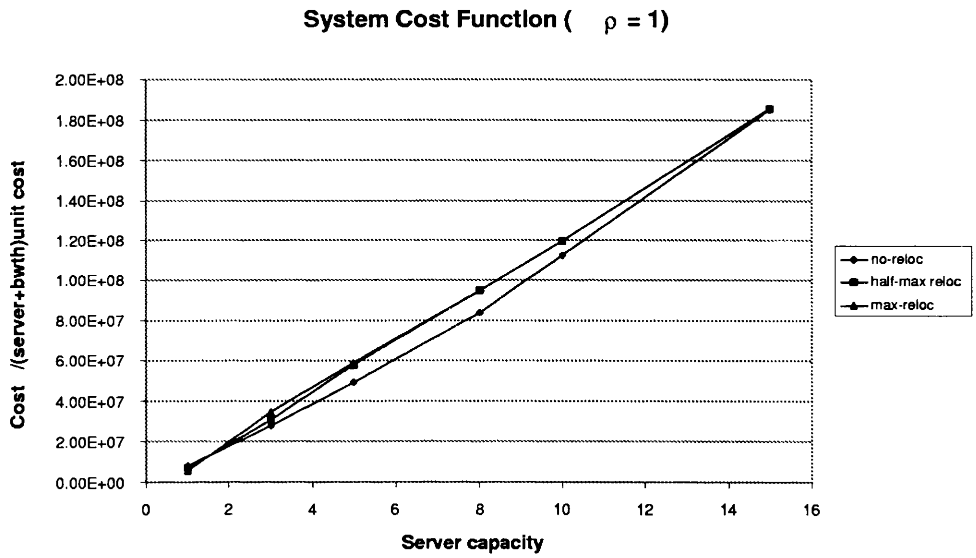

Figure 8.40: System cost function using search distance of 2 hops

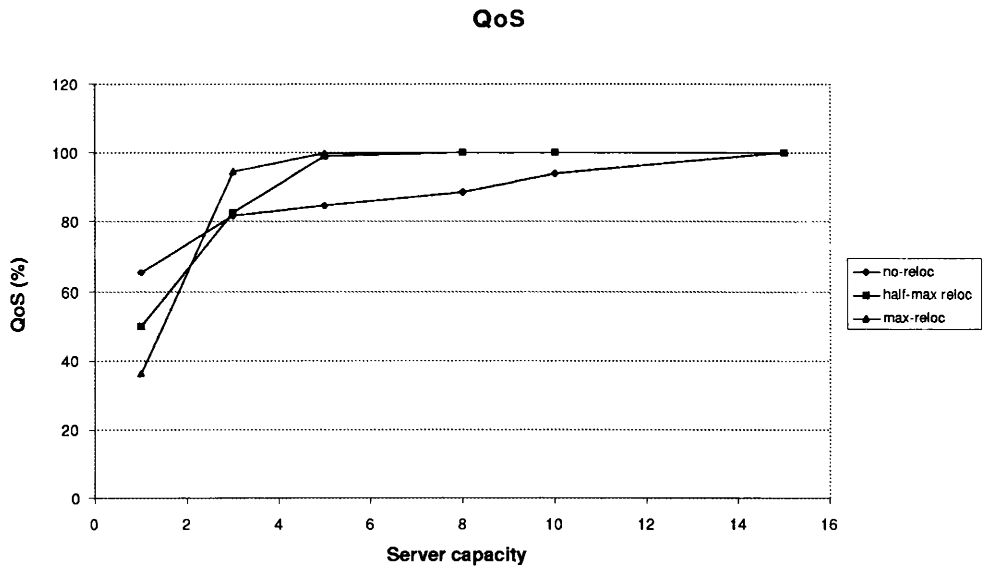

Figure 8.41: QoS achieved using search distance of 2 hops

In the 8-movie local server capacity, system performance is improved. The local cluster capacity is now enough to include all the movies contained in the library. Bandwidth usage drops but not at the same rate as before since we are connecting most users at a maximum distance of 2 hops, see *figure 8.37.* Relocation further improves system performance, since the least popular movies can be reorganised to allow more popular ones to be played in their places; consequently the non-relocating case performs worse (10% less success connection rate), see *figure 8.38.* The number of cut users is due to the high number of least popular movies being set up and the system tendency to decrease bandwidth usage, see *figure 8.39.*

In the 10-movie local server capacity, the cluster capacity is enough to include all the movies contained in the library. Both the relocating scenarios perform the same as the required movie is within the search range of both relocation case, see *figure 8.37.* Their performance is similar to the 1-hop search range case since that was the required capacity to cover the movie library, see *figure 8.38.* There is a number of cut users as the system is able to reorganise movie placement using relocation, see *figure 8.39.* In the non-relocating case cut users exist since the system prefers staring a more popular movie locally in place of one with very few users attached.

In the 15-movie local server capacity, the local server cluster capacity, continues to be enough to include all the movies contained in the library and leave available space as well. Bandwidth usage reduction is small due to the same reason as in the previous case, see *figure 8.37.* Now that most of the least popular movies are played a few fulfil the requirements of cutting down a movie to replace it by a more popular one, see *figure 8.39.*

In the case of using 2 hops search distance, bandwidth usage rises as server capacity increases from 1 to 3 movies, in the cases of half and full range relocation, see *figure 8.38.* The increase is of the order of 2.5 times more bandwidth. In the non-relocation case the bandwidth increase is of the order of 15%. As server capacity increases even further, bandwidth usage falls at a slightly less steep rate. As capacity increases from 3 to five movies the bandwidth usage is reduced by two thirds, in the cases where relocation is used. The non-relocating network uses almost the same amount of bandwidth. As server capacity is increased from 5 to 10 movies, bandwidth usage is further decreased by 20%, in the relocating systems. In the non-relocating system, bandwidth usage is kept at the same level. Finally, as server capacity is increased to its maximum of 15 movies, bandwidth usage decreases further reaching the same levels of the one-movie server capacity cases, in both the relocating systems. In the non-relocating system bandwidth usage is kept at the same level.

The average connection success rate is almost doubled in both relocating cases as server capacity is increased from one to three movies, see *figure 8.37.* In the non-relocating case there is an increase, but is not that steep, of the order of 25%. As server capacity is further increased from 3 to 8 movies, the rate of successful connections is increased by 10%-20%, in all networks. As server capacity is increased to its maximum of 15, successful connection rate is increased but only slightly in the rate of 3%, in all cases. In the relocating cases 100% successful connection rate is reached whereas in the non-relocating case the maximum reached is but with a rather slower rate.

The percentage of cut users is small in all cases. The highest values are encountered in the non-relocating case of 1.2×10^{-4}. The lowest values are encountered in the full-relocating network, see *figure 8.39.*

The system cost rises linearly as server capacity is increased in all cases, see *figure 8.40.* The QoS plot pattern follows the successful connection rate as cut user percentage is low, see *figure 8.41.*

Using 3-hop search distance Here we examine the effects of increasing the local server capacity and using a local search range of 3 hops, for the non-relocating, half-relocating and full-relocating scenario.

As local server capacity is one movie and the search range is increased to 3 hops the local server cluster is capable of accommodating all of the most popular movies of the movie library. The control mechanism gives priority to the very popular movies and tries to set up as many of these as possible. The rest of the movies are started on remote servers and some are relocated later on. Bandwidth usage is increased compared to the 2-hop search range case, see *figure 8.42*. This is due to the increased size of the local server cluster. The user connection rate is increased as well, see *figure 8.43*, since all of the most popular movies are available within the cluster. There are no cut users after the 10 movie server capacity, see *figure 8.44* as only the more popular movies are started and the least popular ones are used for a substantial number of users.

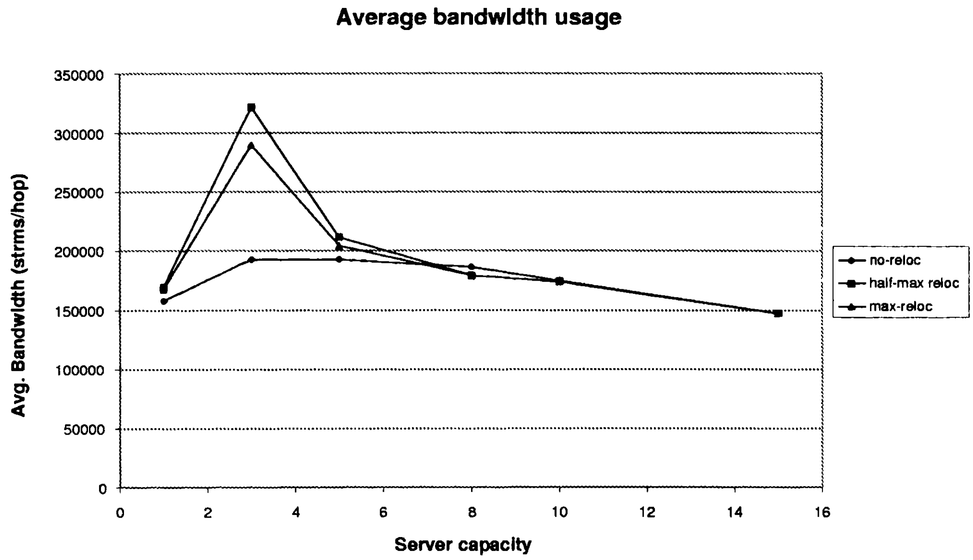

Figure 8.42: Average bandwidth usage using search distance of 3 hops

As local server capacity becomes three movies, the local server cluster is capable of providing all but the last 25 least popular movies in the library. The increased cluster size along with the control mechanism, which is relocating some of the movies in order to utilise better system resources; an increase the bandwidth usage is observed. Consequently the full-relocation case demonstrates a slight increase in bandwidth usage. This increase is less steep, than in the 2-hop search case and so is the gap between the full and half relocation cases, due to the increased server size. On the other hand system performance reaches 90% in all cases as shown in the connection success graph in *figure 8.43*, the increase in connection rate is not as steep, but is similar to that of bandwidth usage. Each server tries to accommodate as many users as possible firstly locally and secondly on the local cluster which in this produces a high success connection rate. There is a percentage of cut users, especially in the non-relocating case, since the increased search range allows the start of less popular movies, but the decreased server capacity does not allow them to continue running, see *figure 8.44*.

As local server capacity increases to 5 movies system performance improves on all counts. Bandwidth usage drops significantly. Now the local server is able to hold 5 of the most popular movies and the local server cluster within 3 hops from it can hold another 195 movies. The total cluster capacity is double the library size. User success rate increases, see *figure 8.43*, since well above 90% of the users can connect within the local cluster. The full relocating scenario performs better, since a small number of users can make way for another more popular movie and due to the increased server size this is now possible is most attempts. In the half relocation case some relocation attempts fail as the required movie is beyond the scope of the search. Compared to the 1-hop search range case, the non-relocating scenario consumes more bandwidth but performs better, see *figure 8.42*, since due to the increased search range more bandwidth is used to access the more popular movies which are now within the user's reach. The half-relocating case consumes the same bandwidth and performs similarly as before, since there are still failed relocation attempts due to the restricted scope of the search range. The full-relocating case now uses less bandwidth and performs very well (95%) since increased cluster capacity and full-relocation allow the placement of the more popular movies to the best access positions for the user pool. In the cut user rate the non-relocating system continues to provide the highest, but reduced due to the increased local server capacity, see *figure 8.44*.

In the 8-movie local server capacity, system performance is improved further. Now the local cluster capacity now 200 movies, is well beyond the size of the

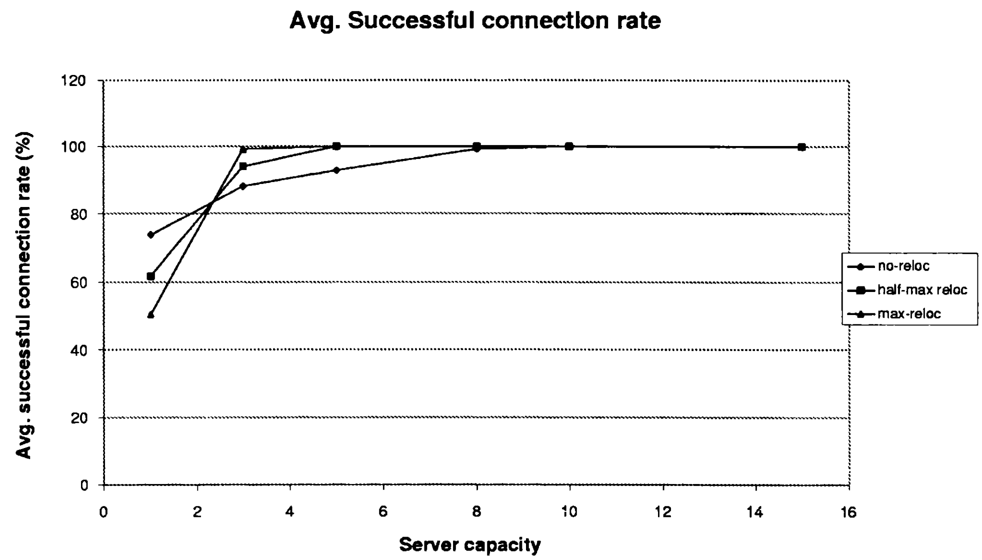

Figure 8.43: Average success connection rate using search distance of 3 hops

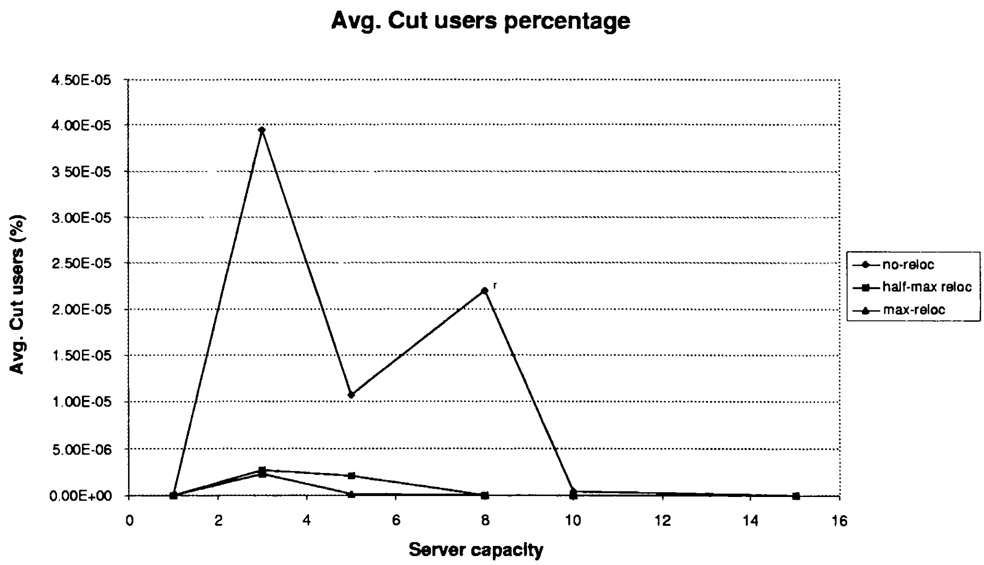

Figure 8.44: Average cut user percentage using search distance of 3 hops

library and enough to accommodate all user requests. Bandwidth usage drops, see *figure 8.42*, but not at the same rate as before since we are connecting most users at a maximum distance of 3 hops. Relocation further improves system performance as shown in *figure 8.43*, since least popular movies can be reorganised to allow more popular ones to be played in their places; and consequently the non-relocating case performs slightly worse (10% less success connection rate). Compared to the 2-hop search range case only the non-relocating scenario shows considerable improvement. Any server size larger than 8, will not produce considerable benefits for the service provider. In the cut user rate the non-relocating system continues to provide the highest rate, slightly increased from the previous case, see *figure 8.44*. The system uses the extra server capacity to place as many popular movies as it can.

In the 10 and 15 local server capacity the system is over engineered, since the local cluster size is 10 to 15 times the library size. Users connect within their local cluster and in the 15-movie case nearer to their local server. The improvement is visible for the non-relocating case, whereas the relocating cases are the same as before, see *figure 8.42*. The cut user rate returns to very low levels since the server space available is enough to accommodate all movies contained in the library, see *figure 8.44*.

In the case of using 3 hops search distance, bandwidth usage rises as server capacity increases from 1 to 3 movies, in the cases of half and full range relocation. The increase is in the order of 2.5 times more bandwidth. In the non-relocation case the bandwidth increase is of the order of 15%. As server capacity increases even further, bandwidth usage falls at a slightly less steep rate, see *figure 8.42*. As capacity increases from 3 to five movies bandwidth usage is reduced by two thirds, in the cases where relocation is used. The non-relocating network uses almost the same amount of bandwidth. As server capacity is increased from 5 to 10 movies, bandwidth usage is further decreased by 20%, in the relocating systems. In the non-relocating system, bandwidth usage is kept at the same level. Finally, as server capacity is increased to its maximum of 15 movies, bandwidth usage decreases further reaching the same levels of the one-movie server capacity cases, in both the relocating systems. In the non-relocating systems bandwidth usage is kept at the same level.

The average connection success rate is almost doubled in both relocating cases as server capacity is increased from one to three movies, see *figure 8.43*. In the non-relocating case there is an increase, but is not that steep, of the order of 10%. As server capacity is further increased from 3 to 8 movies, the rate of successful connections is increased by 10%, in all networks. As server capacity is increased to its maximum of 15, successful connection rate is increased but only slightly in the rate of 3%, in all cases. In the relocating cases 100% successful connection rate is reached whereas in the non-relocating case the maximum reached is but with a slightly slower rate.

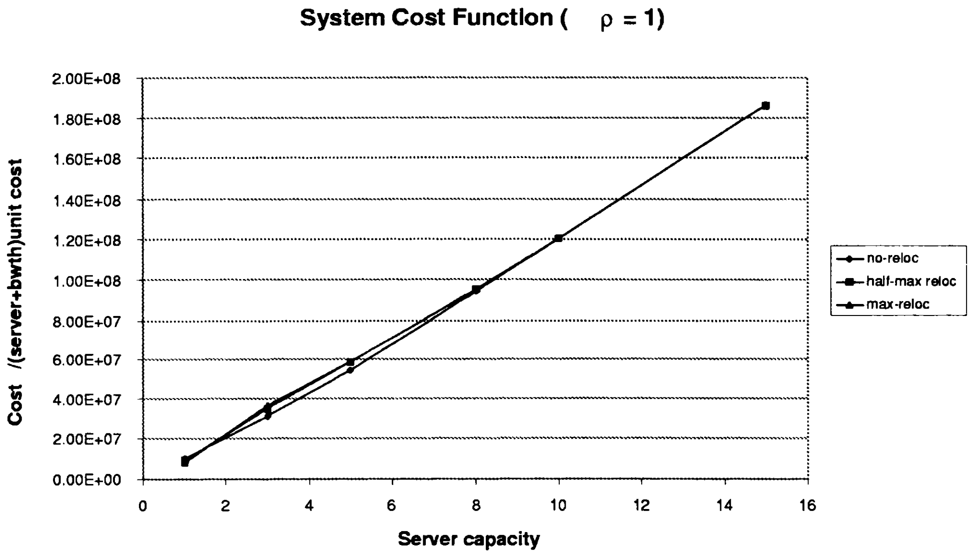

Figure 8.45: System cost function using search distance of 3 hops

The percentage of cut users is small in all cases, see *figure 8.44*. The highest values are encountered in the non-relocating case of 1.2×10^{-4}. The lower values are encountered in the half and full-relocating network of 1×10^{-5} and 5×10^{-6}, respectively, see *figure 8.45*. The system cost rises linearly as server capacity is increased in all cases. The QoS plot pattern follows the successful connection rate as cut user percentage is low, see *figure 8.46*.

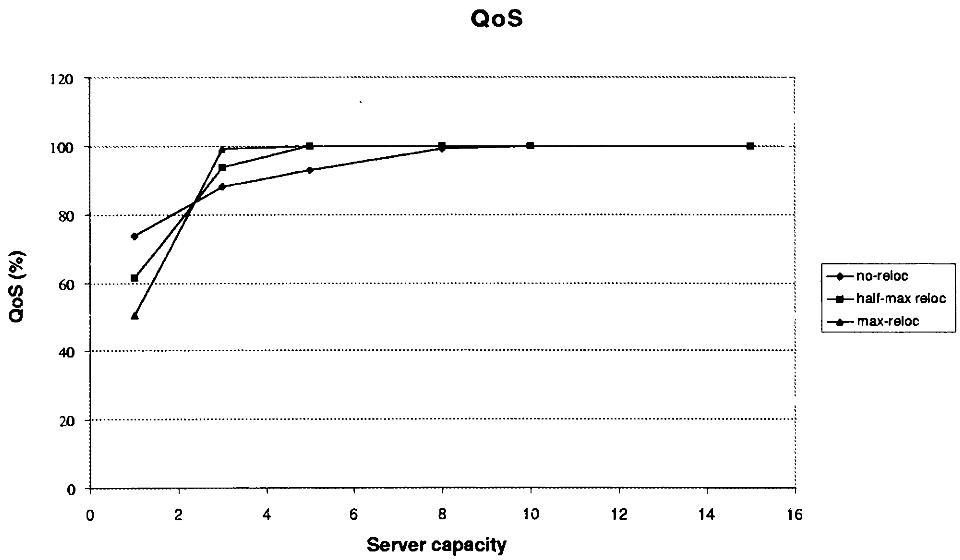

Figure 8.46: QoS achieved using search distance of 3 hops

8.4 Conclusions

The user-access patterns in the multimedia environment are quite complex and unpredictable. The heterogeneity and diversity of IMSs require an efficient, simple and robust control protocol. We demonstrated how the control protocol, can drastically improve network resource usage. We have defined and studied some of the most important control parameters, such as multiplication of material and localisation of connections. These can be used in specifying the carrier network parameters. The proposed control protocol can be used to further probe into the parameter space and study the trade-offs between protocol complexity and network performance.

8.5 Summary

In this chapter we have evaluated the performance of the proposed management system, using different VoD server network scenarios. The management system is based on minimum knowledge of the state of the system, managed to allocate server capacity and bandwidth resources. In the next chapter we compare this performance against two other systems, one completely centralised and one distributed but without any management system.

Chapter 9

Comparison to non-distributed systems

9.1 Introduction

In this chapter we compare the performance of the managed distributed system proposed in the previous chapter to a completely centralised system and to a system of servers without a distributed control mechanism. In the first case we shall show how the core bandwidth requirements increase and in the second we show the deteriorating QoS offered to the end user compared to the managed system with a similar server.

9.2 Distributed servers without control algorithm

In this scenario we used the same network of VoD servers, but without the management system. Every server is capable of providing service only to the users connected to its associated local exchange. The server can not accept or create remote connections either.

The same probability profile and the same user population per local exchange were used. Once a movie was started on a server it could only be changed if the users attached to it were below a certain threshold as well as the playing time was over a certain threshold. The first was done to avoid disconnecting a substantial number of users and the second to give the time margin in order to anticipate possible change in demand.

The number of connected users is given by:

$$U_{at} = \sum_{j}^{t_{ser}} \sum_{i}^{s_{cap}} iz(i) \tag{9.1}$$

where j is the total number of VoD servers in the system, and $z(i)$ the number of users attached to the ith movie as given by the probability function. Only the first i movies in order of popularity, where i is local the server capacity, are played per server. Depending on server size the cumulative user number can range from 10% for 1-movie server capacity to 80% for 15 movies server capacity. This is achieved without using any network bandwidth.

9.3 Centralised system

In the centralised system, only one centralised VoD server is used to provide service. All users from all local exchanges are connected to this one centralised VoD server.

The same probability profile and procedures for starting a new movie and disconnecting one already playing are used. Every connection request is directed to the one centralised VoD server placed on the centre of the square grid network. Bandwidth usage is directed to the server location. Therefore, when dimensioning the network extra capacity should be allowed in this area.

The number of connected users is given by:

$$U_{at} = \sum_{i}^{s_{cap}} iz(i) \tag{9.2}$$

In the centralised video server, only the first i movies in order of popularity, where i is local server capacity, are played. Depending on server size the cumulative user number can range from 10% for 1-movie server capacity to 80% for 15 movies server capacity. The server should be able to support 100 times more video streams. Bandwidth usage is increased as all users are connected to the centralised server

$$B_{at} = \sum_{d}^{s_{net}} c(d) U_{at} d \tag{9.3}$$

where d is the distance in hops from the centralised server, U_{at} is the number of attached users and $c(d)$ is the number of servers situated at distance d from the centralised server.

9.4 Results

The comparison of the performance of the three systems shows the advantages of using the distributed managed system. In terms of bandwidth usage, the non-managed system does not require any inter-server connection and there is no bandwidth usage, but the performance offered, is very poor.

Using 1-hop search distance

Here we compare the performance of the three systems, as the local server capacity increases and the local search distance for the distributed system is 1 hop.

Using 1-hop search distance, bandwidth usage rises as server capacity increases in the centralised system by almost 4 times as server capacity increases, whereas bandwidth usage increases differently in the networked cases. The centralised case is the most bandwidth demanding as expected, since all user traffic is directed to one centralised server, see *figure 9.1.*

Bandwidth usage increases steeply as server capacity increases from 1 to 3 movies, in the cases of half and full range relocation. The increase is of the order of 6 times more bandwidth. In the non-relocation case the bandwidth increase is of the order of 10%. As server capacity increases even further, bandwidth usage falls in a less steep rate. As capacity increases from 3 to five movies the bandwidth usage is reduced by a third, in the cases where relocation is used due to the increased server capacity. The non-relocating network uses almost the same amount of bandwidth. As server capacity is increased from 5 to 10 movies, bandwidth usage is further decreased by one third, in the relocating systems. In the non-relocating system, there is an increase of 50%. Finally, as server capacity is increased to its maximum of 15 movies, bandwidth usage decreases further reaching almost the same levels of the one-movie server capacity cases, in both the relocating systems. In the non-relocating system bandwidth usage is kept at the same level. In the non-networked system there is no bandwidth usage. The centralised case is the most bandwidth demanding as expected. As server capacity increases the number of remote connections increases and traffic is concentrated towards the channels ending on the centralised server.

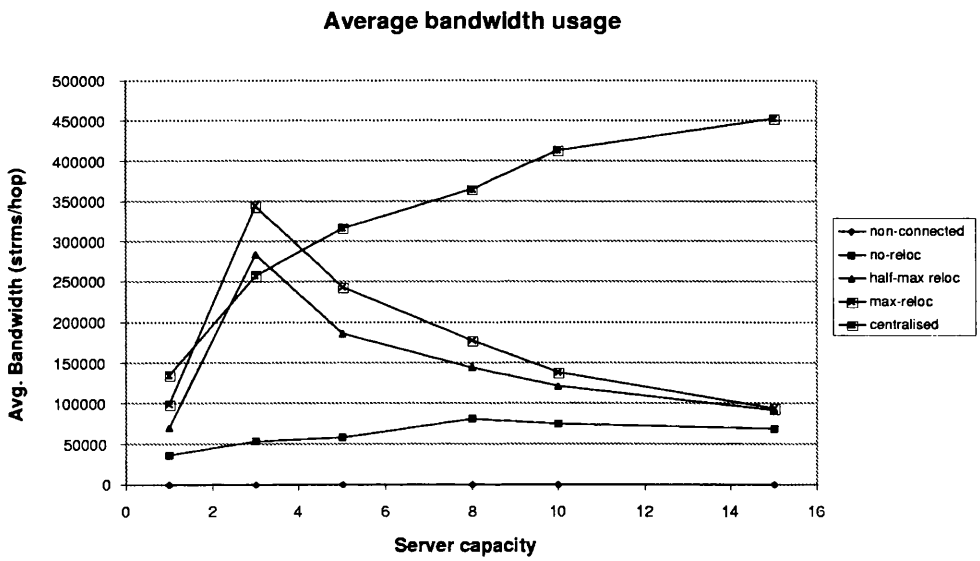

Figure 9.1: Average bandwidth usage using search distance of 1 hop

The average connection rate in both the centralised and non-networked systems is lower than in the networked systems, it rises as server capacity increases by approximately 2.5 times, from the 1-movie capacity case, see *figure 9.2*. In the non-networked system, the servers try to load only the most popular movies and since no inter-working exists they cannot change their movie configuration. In the centralised system the more popular movies are loaded and then cannot be moved due to number of users attached to them. The average connection success rate is almost tripled in both relocating cases as server capacity is increased from one to three movies. In the non-relocating case there is an increase, but is not that steep, it is in the order of 40%. As server capacity is further increased from 3 to 8 movies, the rate of successful connections is increased by 10%-20%, in all networks. As server capacity is increased to its maximum of 15, successful connection rate is increased but only slightly in the rate of 3%, in all cases. In the relocating cases 100% successful connection rate is reached whereas in the non-relocating case the maximum reached is around 80%.

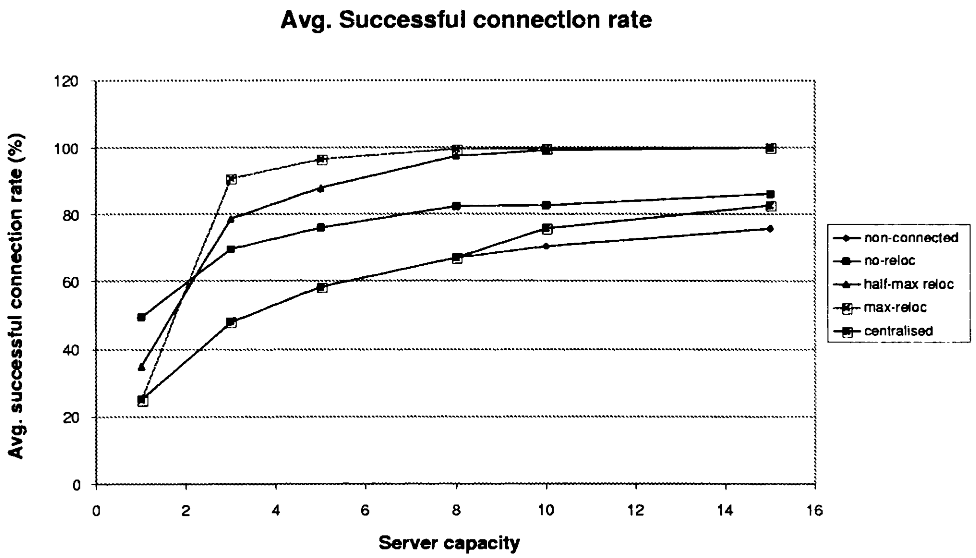

Figure 9.2: Average success connection rate using search distance of 1 hop

The percentage of cut users is quite small in all cases. The highest values are encountered in the non-relocating case of 2.75×10^{-5}. The lowest values are encountered in the full-relocating network. In the centralised and networked systems the cut user rate is almost zero since the criteria for cutting down a movie are never met, due to the absence of server networking, see *figure 9.3*.

The system cost rises linearly as server capacity is increased in all cases and in the non-networked case as well. In the centralised system cost rises from the 10 to 15 server capacity cases by 3 times, see *figure 9.4*. This is due to the increased bandwidth usage.

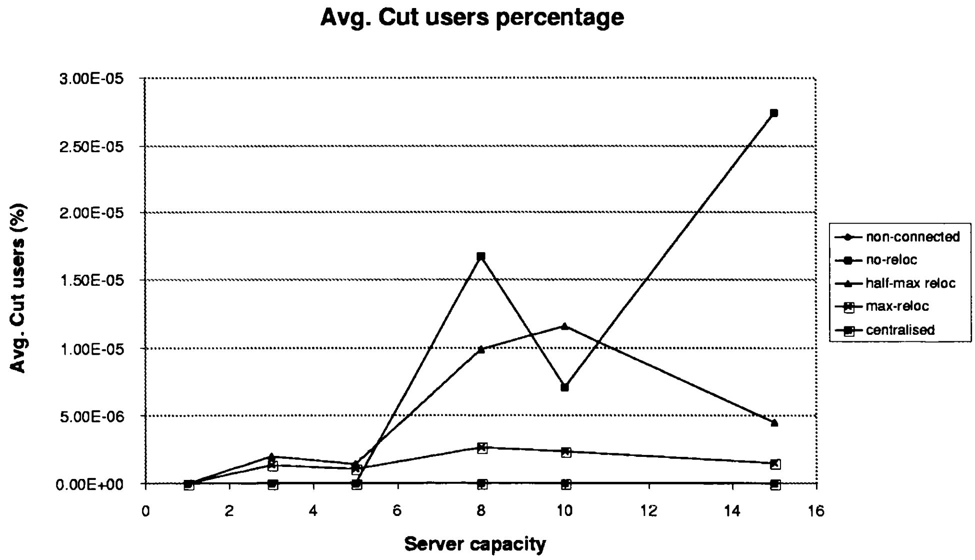

Figure 9.3: Average cut user percentage using search distance of 1 hop

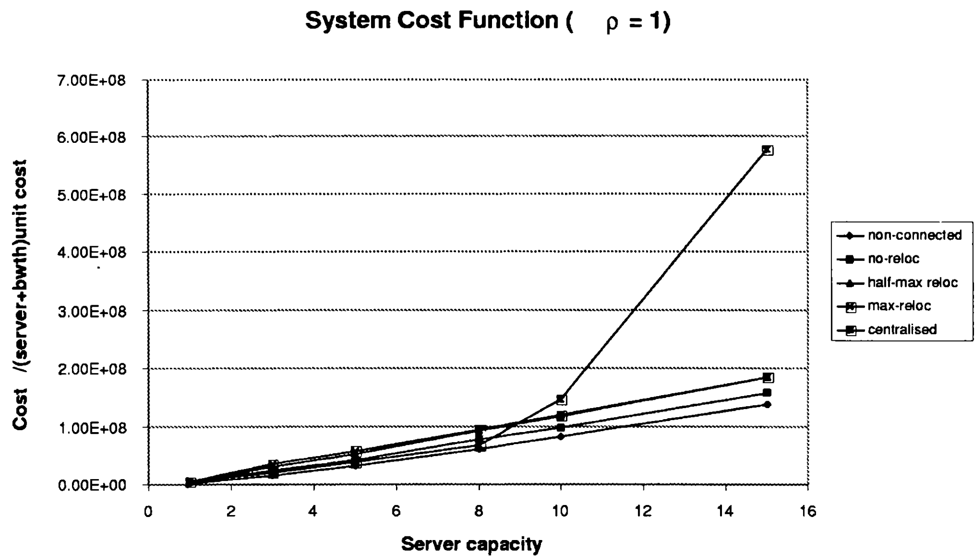

Figure 9.4: System cost function using search distance of 1 hop

The QoS plot pattern follows the successful connection rate as cut user percentage is low and shows the same trends, see *figure 9.5.* The centralised and non-networked systems provide inferior QoS compared to the networked ones, even to the simplest non-relocating one. This is a manifestation of the usefulness of using a distributed system.

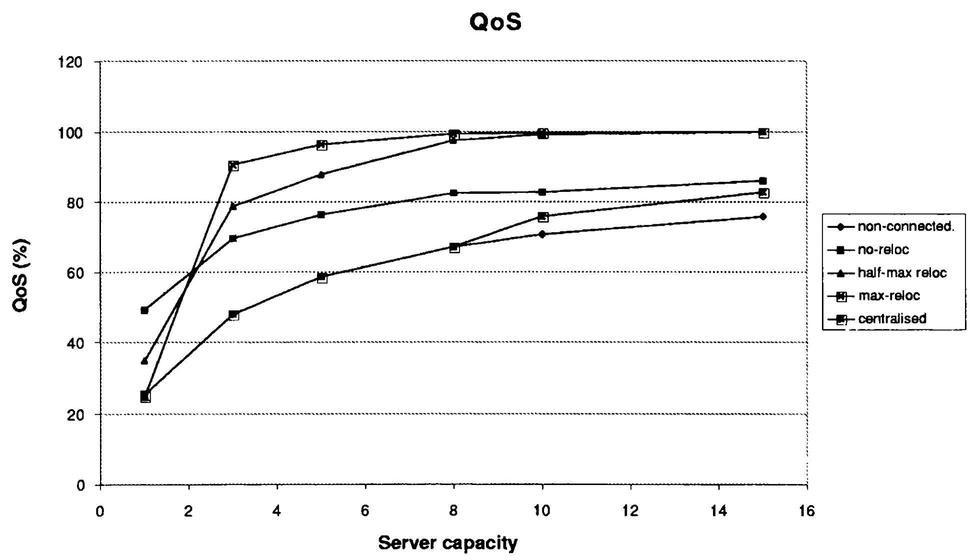

Figure 9.5: QoS achieved using search distance of 1 hop

Using 2-hop search distance

Here we compare the performance of the three systems, as the local server capacity increases and the local search distance for the distributed system is 2 hops.

In the case of using 2 hops search distance, bandwidth usage rises as server capacity increases in the centralised system by almost 4 times as server capacity increases, whereas bandwidth usage increases differently in the networked cases, see *figure 9.6*.

Bandwidth usage rises as server capacity increases from 1 to 3 movies, in the cases of half and full range relocation. The increase is in the order of 2.5 times more bandwidth. In the non-relocating case the bandwidth increase is of the order of 15%. As server capacity increases even further, bandwidth usage falls in a slightly less steep rate. As capacity increases from 3 to five movies the bandwidth usage is reduced by two thirds, in the cases where relocation is used. The non-relocating network uses almost the same amount of bandwidth. As server capacity is increased from 5 to 10 movies, bandwidth usage is further decreased by 20%, in the relocating systems. In the non-relocating system, bandwidth usage is kept at the same level. Finally, as server capacity is increased to its maximum of 15 movies, bandwidth usage decreases further reaching the same levels of the one-movie server capacity cases, in both the relocating systems. In the centralised system bandwidth usage increases following the networked systems up to the 3-movie server capacity. From that point it continues to rise, whereas it falls for the networked systems. This is normal for a centralised system since increase in the number of users accessing the service adds to the total bandwidth used. In the non-relocating system bandwidth usage is kept at the same level.

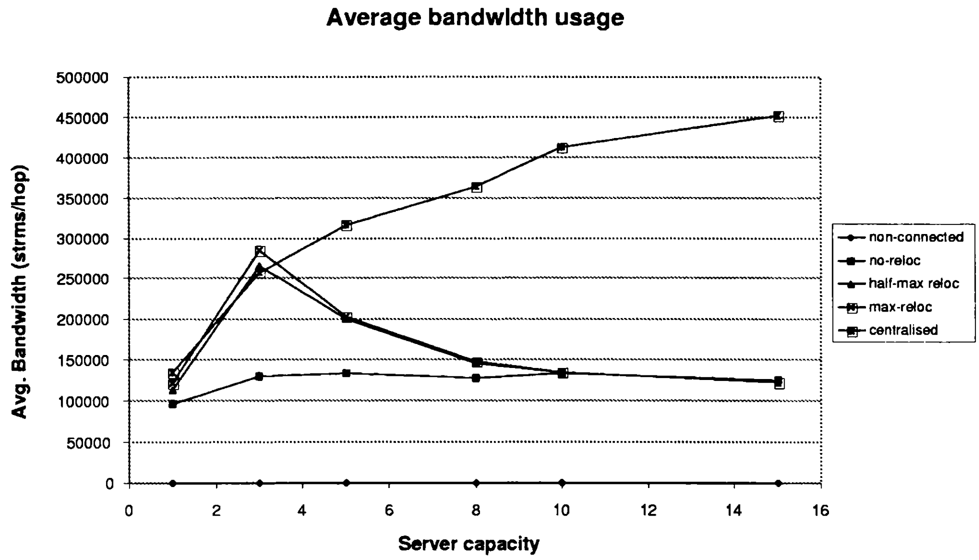

Figure 9.6: Average bandwidth usage using search distance of 2 hops

The average connection rate in both the centralised and non-networked systems is lower than in the networked systems. It rises as server capacity increases by approximately 2.5 times, see *figure 9.7*. The average connection success rate is almost doubled in both relocating cases as server capacity is increased from one to three movies. In the non-relocating case there is an increase, but is not as steep, it is of the order of 25%. As server capacity is further increased from

3 to 8 movies, the rate of successful connections is increased by 10%-20%, in all networks. As server capacity is increased to its maximum of 15, successful connection rate is increased but only slightly in the rate of 3%, in all cases. In the relocating cases 100% successful connection rate is reached whereas in the non-relocating case the maximum is reached but with a rather slower rate. The centralised and non-networked systems provide an average connection rate reduced on average by more than 20% compared to the networked systems. This is due to their inability to use their resources in cooperation. The servers provide only the most popular movies and are incapable of sharing their resources.

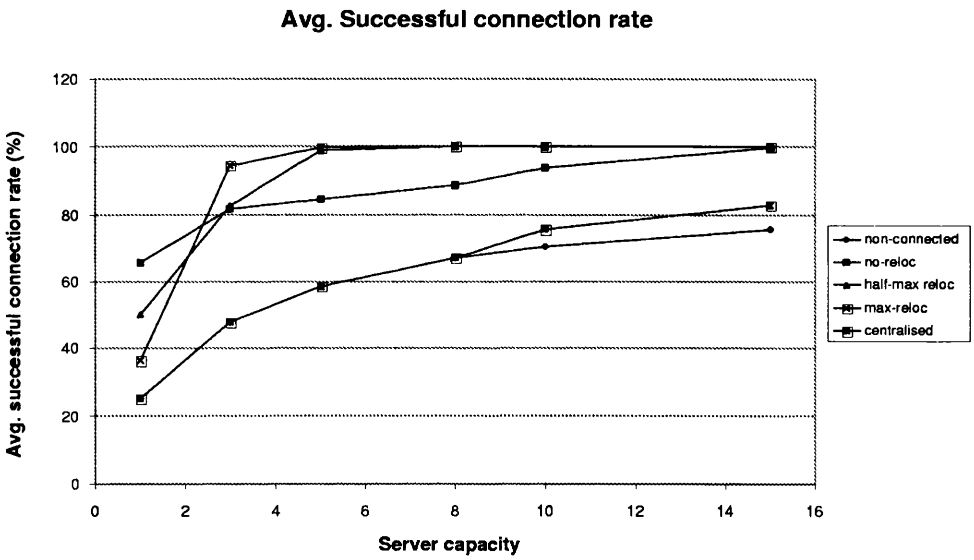

Figure 9.7: Average success connection rate using search distance of 2 hops

The percentage of cut users is small in all cases. The centralised and non-networked systems never meet the criteria for cutting down a movie. The highest values are encountered in the non-relocating case on the 3 movie server capacity of 1.2×10^{-4}. The other two networked systems produce a smaller cut user percentage of 1.5×10^{-5}. The lowest values are encountered in the full-relocating network. In the centralised and networked systems the cut user rate is almost zero, see *figure 9.8*. Relocation and server reconfiguration are allowed at the cost of slightly reducing QoS.

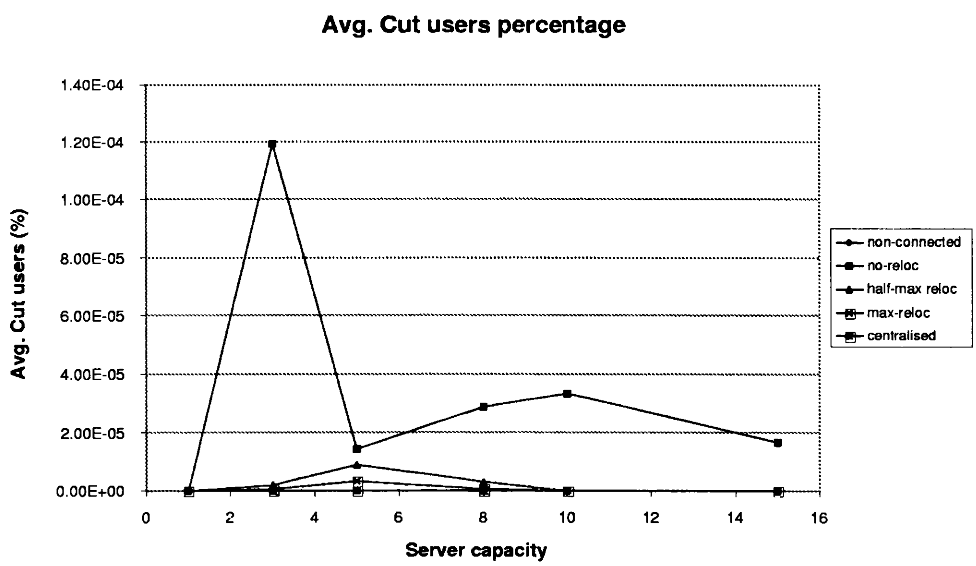

Figure 9.8: Average cut user percentage using search distance of 2 hops

System cost rises linearly as server capacity is increased in all cases. In the centralised system cost rises from the 10 to 15 server capacity cases by 3 times, see *figure 9.9.* This is due to the considerable increase in bandwidth in the centralised system as server capacity rises.

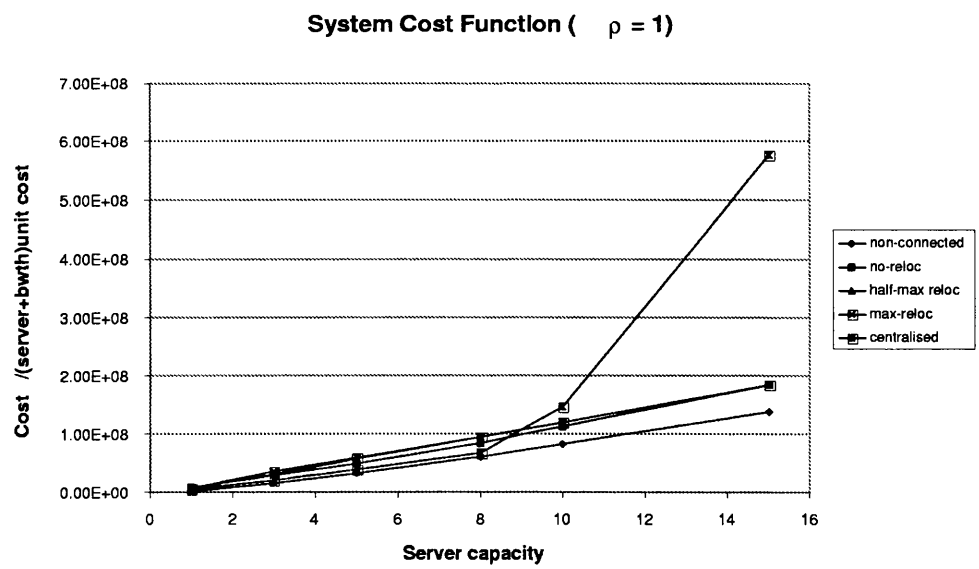

Figure 9.9: System cost function using search distance of 2 hops

The QoS plot pattern follows the successful connection rate as cut user percentage is low, see *figure 9.10.* The non-networked systems produce a 20% reduced QoS due to the high number of users that are not granted their request.

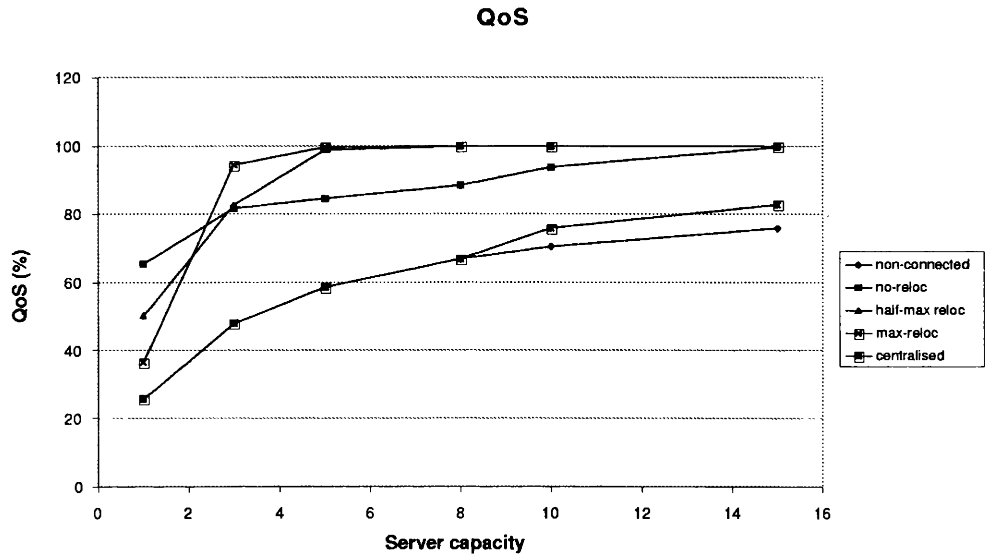

Figure 9.10: QoS achieved using search distance of 2 hops

Using 3-hop search distance

Here we compare the performance of the three systems, as the local server capacity increases and the local search distance for the distributed system is 3 hops.

In the case of using 3 hops search distance, bandwidth usage rises as server capacity increases in the centralised system by almost 4 times as server capacity increases, whereas bandwidth usage increases differently in networked cases, *figure 9.11.*

Bandwidth usage rises as server capacity increases from 1 to 3 movies, in the cases of half and full range relocation. The increase is of the order of 2.5 times more bandwidth. In the non-relocation case, the bandwidth increase is of the order of 15%. As server capacity increases even further, bandwidth usage falls in a slightly less steep rate. As capacity increases from 3 to five movies bandwidth usage is reduced by two thirds, in the cases where relocation is used. The non-relocating network uses almost the same amount of bandwidth. As server capacity is increased from 5 to 10 movies, bandwidth usage is further decreased by 20%, in the relocating systems. In the non-relocating system, bandwidth usage is kept at the same level. Finally, as server capacity is increased to its maximum of 15 movies, bandwidth usage decreases further reaching the same levels of the one-movie server capacity cases, in both the relocating systems. In the centralised system bandwidth usage continues to rise as more and more distant users are connected to the one centralised server. In the non-relocating system bandwidth usage is kept at the same level.

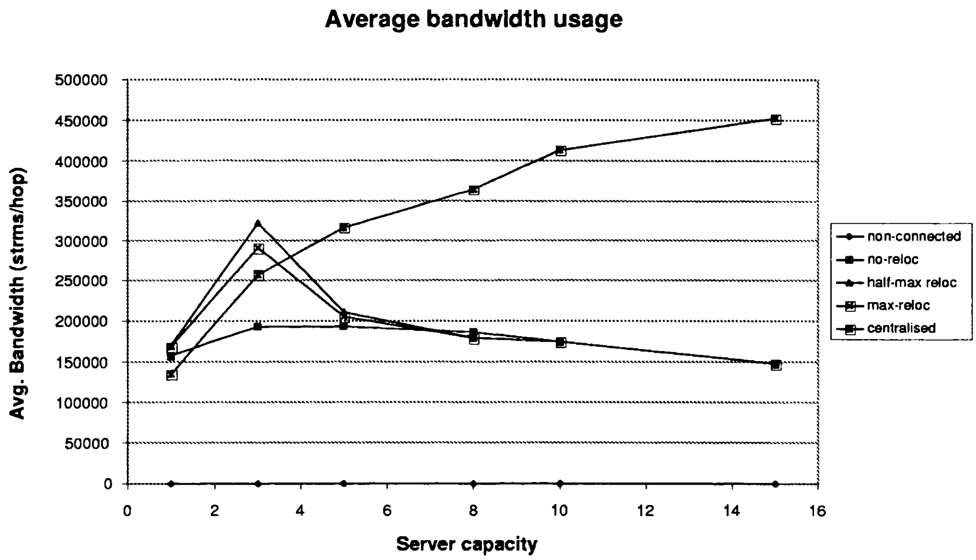

Figure 9.11: Average bandwidth usage using search distance of 3 hops

The average connection rate in both the centralised and non-networked systems is lower than in the networked systems. It rises as server capacity increases by approximately 2.5 times, see *figure 9.12*. The average connection success rate is almost doubled in both relocating cases as server capacity is increased from one to three movies. In the non-relocating case there is an increase, but is not that steep, of the order of 25%. As server capacity is further increased from 3 to 8 movies, the rate of successful connections is increased by 10%-20%, in all networks. As server capacity is increased to its maximum of 15, successful connection rate is increased but only slightly in the rate of 3%, in all cases. In the relocating cases 100% successful connection rate is reached whereas in the non-relocating case the maximum is reached but with a rather slower rate. The non-networked systems behave as before but now the difference to the networked ones increases further.

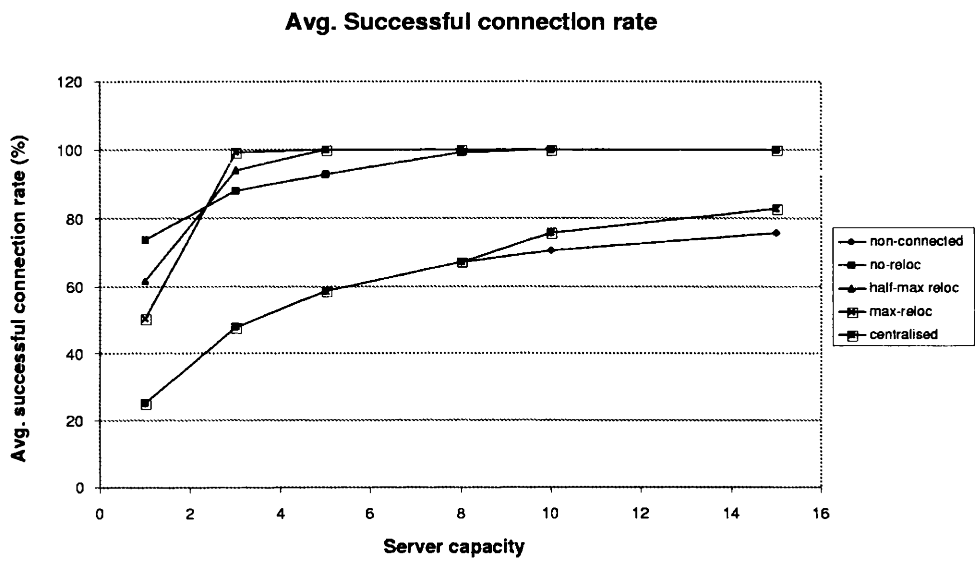

Figure 9.12: Average success connection rate using search distance of 3 hops

The percentage of cut users is quite small in all cases, see *figure 9.13.* The non-networked systems still do not produce any cut users. The highest values are encountered in the non-relocating case of 4.00×10^{-5} at the 3-movie server capacity network. The lowest values are encountered in the full-relocating network.

The system cost rises linearly as server capacity is increased in all cases. In the centralised system cost rises from the 10 to 15 server capacity cases by 3 times, see *figure 9.14.* Again here the cost of using excessive bandwidth gives a rise to the total cost of the centralised system.

The QoS plot pattern follows the successful connection rate as cut user percentage is low, see *figure 9.15.* The relocating systems perform almost perfectly,

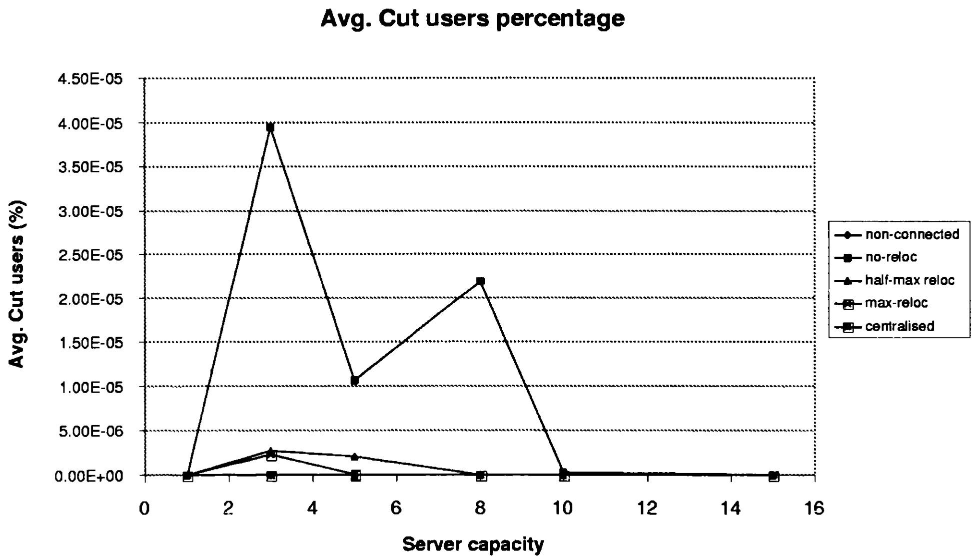

Figure 9.13: Average cut user percentage using search distance of 3 hops

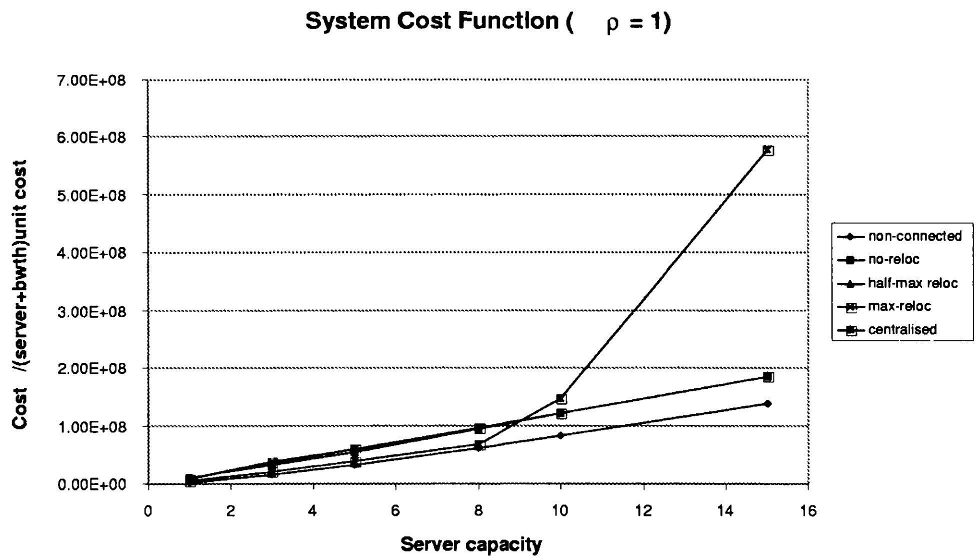

Figure 9.14: System cost function using search distance of 3 hops

whereas the non-networked ones perform on average 30% worse. Relocation and networking exploit much better system resources.

The performance and cost of the centralised system are considerably higher than in the networked systems. The non-networked system is comparable in

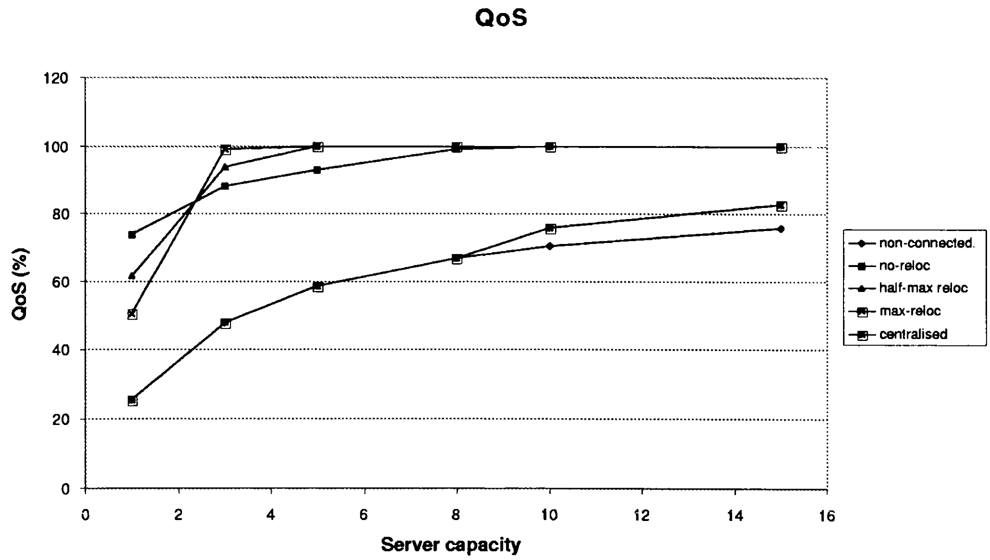

Figure 9.15: QoS achieved using search distance of 3 hops

cost with the networked systems but lacks in performance. The networked solution provides 20%-30% better performance for the same cost.

9.5 Conclusion

Comparing the performance of the distributed control system to a control-less and to a centralised one, gives a good insight as to why such a system should use distributed control. In both cases QoS service deteriorates drastically when the aggregate server capacity is below the library size. Because of the lack of flexibility in both control-less and centralised systems users are denied service or are even being disconnected due to the inefficiency of the system to use its resources.

In the control-less system poor use of the available VoD servers is achieved, as almost all the demand for the rest of the movies in the library that cannot be accommodated in the local server, are rejected. Those that are served are at the expense of deteriorating QoS as the users of a playing movie have to be disconnected in order for the new movie to start.

In the centralised scenario, when the centralised server capacity does not exceed the size of the library, requests are served at the expense of over using core network bandwidth. In the case where the server capacity is superseded by library size, we have additional deterioration of service due to the rejected users.

9.6 Summary

In this section we have compared our proposed VoD service design to two alternative scenarios. One scenario was that of a completely localised structure and the other of a completely centralised one. In both cases the advantages of the distributed system were proved both in the QoS and system resources usage aspects.

Chapter 10

Conclusions

10.1 Introduction

In this thesis, a design for a distributed VoD service has been presented, concentrating on two aspects: the VoD server; and the server network management system. Emphasis has been placed on the simplicity, scalability and performance of the design.

The service enabling technologies and commercial VoD trials are reviewed, there we present our proposed VoD server design. The server design is based on short interval staggered multicast and two-tier user stream flow control. The design is scalable as more units can be added as required.

A review of other distributed service systems is followed by the presentation of our distributed service management system. Movie and connection placements are carried out based on movie popularity. Popular movies are replicated across more servers, and placed close to the end user, to minimise core network bandwidth usage. The least popular movies are provided when and wherever server space is available. Distant connections to these movies are allowed since only a small number of users require them.

Additionally, less popular movies are relocated to other servers to provide server space for other more popular ones. Finally, least popular movies are terminated to allow a more popular movie to start.

We tested our proposed control system on a 10×10 square grid network of VoD servers of varying movie capacity from 1 to 15 movies. The library size was 100 movies, and was kept constant. We used a non-relocating and two relocating algorithms. Of the relocating algorithms the first could relocate a movie to an area of radius of half the maximum network distance and the second to full. We monitored bandwidth usage, connection success rate, cut user rate, cost and QoS. The relocating systems performed better for server capacity of more than 1 movie. Full relocation generally performed better than half depending on server capacity, especially at lower server capacities. As server capacity increased both relocating algorithms performed similarly.

Finally, we compared our system to a non-interconnected network of servers and to a centralised system. The system proposed and studied has performed better and provided improved QoS.

10.2 Contributions of this thesis

The main points and conclusions derived from the present work can be summarised as follows:

- In the fourth chapter: The most important VoD server characteristics are identified. These parameters need to be included in all VoD server designs, to improve efficiency and performance. The VoD server design is presented and analysed. This design is scalable and can provide interactive control to the end user. Depending on the dimensioning of its parameters it can be used from multicast to true-VoD. This design can form the basis of future commercial service implementations.

- In the sixth chapter: The most important network resources are identified. Bandwidth usage and server capacity are the two most important parameters in the VoD service network and play a key role in its function and performance. Their management forms the basis of the distributed management system.

- In the eighth chapter: We study the performance of our proposed service design under varying network and control parameters. Both relocating systems perform better for medium server capacities, since the network has enough server capacity to accommodate relocation. The improved performance of the relocating networks continues to higher servers capacities. These results prove the usefulness of relocation on our networked system. Our system works by forming local server clusters, which decentralise bandwidth usage and provide improved movie availability to the end users. The addition of relocation improves performance even more.

- In the ninth chapter: We compare our distributed VoD server network to two non-networked systems. One is a centralised one and the other a non-interconnected system of video servers. Our networked systems performed better than the non-networked ones further proving the value of distributed control.

10.3 Suggestions for further work

Further work based on this study could include:

- Introducing a more complex management system based on more network status information and optimisation methods to compare performance and cost.

- Applying this method to multimedia services where read-write processes are performed.

10.4 Concluding Remarks

In this thesis, a design for a VoD server and a design for a distributed server managing system, the two most important VoD service components, were presented. The management system was evaluated for various sets of network and management system parameters. Finally, its performance was compared against a centralised and non-networked system. The advantages of distribution were shown.

Bibliography

[1] A .D .Gelman, H .Kobrinski, L. S. Smoot, S. B. Weinstein, M. Fortier and D. Lemay. A store-and-forward architecture for video-on-demand service. *Canadian Journal of Electronic and Computer Engineering*, 18(1):37–40, 1993.

[2] John Yanosy and Brett Leida. Content Hosting in Full Service Broadband Delivery infrastructures: Efficiencies, Architecture and Policy . In *The 23rd Annual Telecommunications Policy Research Conference Solomons, Maryland*, 1995.

[3] Scott Barnett and Gary Anido. A cost comparison of Distributed and Centralised Approaches to Video-on-Demand. *IEEE Journal on Selected Areas in Communications*, 14(6):1173–1183, August 1996.

[4] Alexander D. Gelman and Lanny S. Smoot. An architecture for interactive applications. In *IEEE International Conference on Communications 1993.*

[5] Bur Goode. Scanning the issue. *Proc. of the IEEE*, 85(12), December 1997.

[6] Independent JPEG Group. *jpeg-info@uunet.uu.net.*

[7] Masahisa Kawashima, Cheng-tie Chen, Fure-Ching Jeng and Sharad Singhal. Adaption of the MPEG Video-Coding Algorithm to Network Applications. *IEEE Journal on Selected Areas in Communications*, 3(4):261–269, August 1993.

[8] Pramod Pancha and Magda El Zarki. MPEG Coding for Variable Bit Rate Video Transmission.

[9] Soung C. Liew and Chi yin Tse. Video Aggregation: Adapting Video Traffic for Transport over Broadband Networks by Integrating Data Compression and Statistical Multiplexing. *IEEE Journal on Selected Areas in Communications*, 14(6):1123–1137, August 1996.

[10] http://www.vela.com/ arun/COMP.html.

[11] N. G. Cole. Asymmetrical digital line technology – A basic oveview. *Br. Telecom. Technol. J.*, 12(1):118–125, January 1994.

[12] D. L. Waring, D. S. Wilson and T. R. Hsing. Fibre upgrade strategies using high-bit-rate copper technologies for video delivery. *IEEE J. Lightwave Technol.*, 10(11):1743–1750, November 1992.

[13] Walter Y. Chen and David L. Waring. Applicability of ADSL to support Video Dial Tone in the Copper Loop. *IEEE Communications Magazine*, pages 102–109, May 1994.

[14] Maurice Gagnaire. An Overview of Broad-Band Access Technologies. *Proceedings of the IEEE*, 85(12):1958–1972, December 1997.

[15] J. R. Fox, J. Marshall and G. Mulholland. Lessons learned from the Bishop's Strotford optical trial. In *ISSLS '93 Conference in Vancouver*, 27 September – 1 October 1993.

[16] R. D. Carver. Millimetre-wave radio for broadband access. In *ICC 91, Denver, CO, USA*, 23-26 June 1991.

[17] http://www.isdn.bt.com/.

[18] *IEEE Journal on Selected Areas in Communications*, 15(6), August 1997.

[19] Chi-Yuan Hsu, Antonio Ortega and Amy R. Reinbman. Joint Selection of Source and Channel Rate for VBR Transmission Under ATM Policing Constraints. *IEEE Journal on Selected Areas in Communications*, 15(6):1016–1028, August 1997.

[20] Wenjun Luo and Magda El Zarki. Quality Control for VBR Video over ATM Networks. *IEEE Journal on Selected Areas in Communications*, 15(6):1029–1039, August 1997.

[21] Maher Hamdi, James W. Roberts and Pierre Rolin. Rate Control for Video Coders in Broad-Band Networks. *IEEE Journal on Selected Areas in Communications*, 15(6):1040–1051, August 1997.

[22] ATM Forum Technical Working Group. SAA Audio-visual multimedia service (AMS), iplementation agreement. *Contribution 95–0012, Burlingame, CA*, Feb. 1995.

[23] MPEG-2 Systems Committee. MPEG-2 systems working draft. *ISO/IEC/JTC1/SC29/WG-11-N0501.*

[24] Jean M. McManus and Keith W. Ross. Video-on-Demand Over ATM: Constant-Rate Transmission and Transport. *IEEE Journal on Selected Areas in Communications*, 14(6):1087–1098, August 1996.

[25] W. Verbiest, L. Pinnoo and B. Voeten. The impact of ATM concept on video coding. *IEEE Journal on Selected Areas in Communications*, 6(9):1623–1632, August 1988.

[26] D. G. Morrison. Variable bit-rate coding for asynchronous tranfer mode networks. *Br. Telocom. Technol. J.*, 8(3):70–80, July 1990.

[27] D. Reininger and D. Raychaudhuri. Bit-rate characteristics of a VBR MPEG encoder for ATM networks. In *IEEE ICC'93, Geneva, Switz*, pages 517–521, May 1993.

[28] B. Melamed B. Sengupta D. Reininer, D. Raychaudhuri and J. Hill. Statistical Multiplexing of VBR MPEG compressed video on ATM networks. In *INFOCOM'93, San Francisco, CA, USA*, pages 919–926, March 1993.

[29] F. Kashino et al. Variable bitrate coding of video signals for ATM networks. *IEEE Journal on Selected Areas in Communications*, 7(5):801–806, June 1989.

[30] Vahe Balabanian, Liam Casey and Nancy Greene. Digital Storage of Media-Command and Control Protocol Applied to ATM . *IEEE Journal on Selected Areas in Communications*, 14(6):1162–1172, August 1996.

[31] Information Technology. Generic Coding of Moving Pictures and Associated Audio: Systems ITU-T Recommendation H.222.0/ISO/IEC 13818-1 International Standard. *(MPEG-2 Systems), ISO/IEC/JTC1/SC29/WG11, N0801rev.*, 1995.

[32] Information Technology. Generic Coding of Moving Pictures and Associated Audio: Digital Media Command and Control-ISO/IEC 13818-6 Draft International Standard. *ISO/IEC JTC1/SC29/WG11 MPEG95/N1100*, Nov. 1995 1995.

[33] A. Mankin S. Bradner. RFC 1752, The Recommendation for the IP Next Generation Protocol. January 1995.

[34] http://www.6bone.net/ngtrans.html.

[35] Maurizio Decina and Vittorio Trecordi. Convergence of Telecommunications and Computing to Networking Models for Integrated Services and Applications. *Proceedings of the IEEE*, 85(12):1887–1914, December 1997.

[36] Mitchel Mandjes Kees Van Der Wal and Harrie Bastiaanseen. Delay Performance Analysis of the New Internet Services with Guaranteed QoS. *Proceedings of the IEEE*, 85(12):1947–1957, December 1997.

[37] Paul P. White and Jon Crowcroft. The Integrated Services in the Internet: State of the Art. *Proceedings of the IEEE*, 85(12):1934–1946, December 1997.

[38] David Brierly and Nicholas Moss. Telekom's regulator shows (a little) spine. *The European*, page 18, 4 January - 14 January 1998.

[39] http://www.oftel.gov.uk/.

[40] Dirk De Schoenmacker Bruno Suard, Leopold Verbist. Update on vod trilas. *International Conference on Communication Technology*, 2:1033–1036, 1996.

[41] Larry Plumb and Ginger Fisk. Bell Atlantic services' VOD Market Trial Results. *Bell Atlantic Media Relations Home Page.*

[42] A. Dan, D. Sitaram, P. Shahabuddin. Sceduling policies for an on demand video server with batching. *IBM Research Report RC 19381.*

[43] P. Samuelson. Copyright's Fair Use Doctorine and Digital Data. *Comm. ACM*, 37(1):21–27, January 1994.

[44] D. Jadav, A. Choudhary and P. B. Berra. An evaluation of design trade-offs in a hifg-performance media-on-demand server. *Multimedia Systems*, 5(1):53–68, 1997.

[45] Cyril U. Orji, Patrick O. Bobbie and Kingsley C. Nwosu. Spatio-temporal effects of multimedia objects storage and delivery for video-on-demand systems. *Multimedia Systems*, 5(1):39–52, 1997.

[46] A. Srivastava, A. Kumar and A. Singru. Design and analysis of a video-on-demand server. *Multimedia Systems*, 5(1):238–254, 1997.

[47] W. D. Sincoskie. System architecture for a large scale video on demand service. *Computer Network and ISDN Systems*, 22:155–162, 1991.

[48] Herbert Taub and Donald L. Schilling. *Principles of communication systems.* McCraw-Hill, 1986.

[49] J. R. Allen, B. L. Heltai, A. H. Koenig, D. F. Snow and J. R. Watson. VCTV: A Video-On-Demand Market Test. *AT & T Technical Journal*, 72(1):7–14, January/February 1993.

[50] Gordon Kerr. A review of fully interactive video on demand. *Signal Processing: Image Communication*, (8):173–190, 1996.

[51] http://www.cray.com/products/systems/. Cray Systems.

[52] http://www.ssd.intel.com/.

[53] Editor:Sape Mullender. *Distributed Systems.* Addison-Wesley, 1993.

[54] Stefano Ceri and Giuseppe Pelagatti. *Distributed Databases : principles and systems.* New York : McGraw-Hill, 1984.

[55] Alan Hutchinson. *Algorithmic Learning.* Oxford : Clarendon Press, 1994.

[56] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, 1988.

[57] L. C. Wolf, C. Griwodz and R. Steinmetz. Multimedia Communication. *Proceedings of the IEEE*, 85(8):1915–1933, December 1997.

[58] F. M. Renon and M. Ullio. Interactive Multimedia Services: Hierarchical vs. Full Meshed Network Architectures. In *Network and Optical Communication II (NOC' 96)*, pages pp.75–82, 1996.

[59] J. R. Allen B. L. Heltai A. H. Koening D. F. Snow and J. R. Watson. VCTV: A Video–on–Demand market test. *AT & T Technical Journal*, 72(1):pp. 7–14, Jan./Feb. 1993.

[60] K. Papanikolaou and M. Wilby. Design outline for Video–on–Demand server. In *Fifth IEE Conference on Telecommunications*, pages pp. 313–317, 1995.

[61] K. Papanikolaou and M. Wilby. Management Scheme for VoD Server Network . In *Network and Optical Communications 1996 (NOC'96, Heidelberg): ATM Networks and LAN's*, pages pp. 206–213, 1996.

[62] Y. S. Chen. Mathematical modelling of empirical laws in computer application: A case study. *Comput. Math. Applicat.*, pages 77–78, Oct. 1992.

[63] J. Nussbaummer B. V. Patel F. Schaffa and P. G. Sterbenz. Networking Requiremnts for Interactive Video on Demand. *IEEE Journal on Selected Areas in Telecommunications*, 13(5):pp. 779–787, June 1995.